DOCUMENT RESUME

ED 092 508 95 SP 008 097

AUTHOR Gropper, George L.
TITLE A Technology for Developing Instructional Materials. Vol. 3, Handbook. Part F, Develop Diagnostic and Evaluative Tests.
INSTITUTION American Institutes for Research in the Behavioral Sciences, Pittsburgh, Pa.
SPONS AGENCY Office of Education (DHEW), Washington, D.C.
PUB DATE Mar 73
CONTRACT OEC-0-70-4776(520)
NOTE 159p.; For related documents, see SP 008 090-096 and 098-104

EDRS PRICE MF-$0.75 HC-$7.80 PLUS POSTAGE
DESCRIPTORS *Assignments; Development; *Diagnostic Tests; Educational Development; Educational Research; *Evaluation Techniques; *Guides; Instructional Materials; Manuals; Personnel; *Tests

ABSTRACT

This document, sixth in a series of 11 subvolumes of a handbook prepared to provide training for educational research and development personnel in the development of instructional materials, deals with the task of developing diagnostic and evaluative tests. The document is organized according to the four sequential steps involved in performing the task. The four steps involve: (a) planning the development of tests; (b) developing tests for proficiency in criterion behavior which is taught in each instructional unit; (c) developing diagnostic tests to identify the reason for failure to acquire the criterion behavior taught in each instructional unit; and (d) trying out and revising testing procedures. More specific substeps list procedures for performing each step. (PD)

AIR-20241-9/71-RP

A Technology For Developing Instructional Materials

3 HANDBOOK

ED 092508

A. PLAN STUDY OF CRITERION BEHAVIORS

B. COLLECT AND ANALYZE DATA ABOUT CRITERION BEHAVIORS

C. SEQUENCE AND GROUP CRITERION BEHAVIORS

D. STATE CRITERION AND PREPARATORY OBJECTIVES

E. PLAN SIMULATION BASED ON INSTRUCTIONAL AND LOGISTICAL NEEDS

F. DEVELOP DIAGNOSTIC AND EVALUATIVE TESTS

G. FORMULATE INSTRUCTIONAL STRATEGIES

H. PLAN ACCOMMODATION OF INDIVIDUAL DIFFERENCES

I. DEVELOP INSTRUCTIONAL MATERIALS

J. EVALUATE INSTRUCTIONAL MATERIALS

X. INDEX

AUTHOR:

George L. Gropper

Published by:

AMERICAN INSTITUTES
FOR RESEARCH
Pittsburgh, Pennsylvania

© Copyright
MARCH, 1973

SP008 091

U.S. DEPARTMENT OF HEALTH
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY.

PERMISSION TO REPRODUCE THIS COPYRIGHTED MATERIAL HAS BEEN GRANTED BY
George L. Gropper
TO ERIC AND ORGANIZATIONS OPERATING UNDER AGREEMENTS WITH THE NATIONAL INSTITUTE OF EDUCATION. FURTHER REPRODUCTION OUTSIDE THE ERIC SYSTEM REQUIRES PERMISSION OF THE COPYRIGHT OWNER.

VOLUMES IN THIS SERIES

1. USER'S MANUAL
2. ORIENTATION
3. HANDBOOK
   (eleven sub-volumes)
4. WORKBOOK
5. FINAL EXERCISES

FOREWORD

This is one of a series of eleven HANDBOOK sub-volumes which has been prepared to provide training for educational R&D personnel in the development of instructional materials.

The USER'S MANUAL, which accompanies the series, describes the role each volume is designed to play and the sequence recommended for its use in the training process. The user is, therefore, urged to read the instructions in the USER'S MANUAL before using this or any other separate volume.

ACKNOWLEDGMENTS

The materials in this volume were prepared under a contract from the U.S. Office of Education, Contract No. OEC-0-70-4776(520). Dr. George L. Gropper, Director of Instructional Media Studies, served as principal investigator.

U.S.O.E. sponsorship does not in any way imply official endorsement of the views expressed in this volume.

The author is indebted: to Dr. Robert Fitzpatrick for reviewing portions of the series of volumes and for informal discussions concerning several training issues; to Mrs. Zita Glasgow for the first and critical use of this volume; and, not least, to Miss Kathleen Gubala for her tireless preparation of the complex manuscript required by this HANDBOOK.

George L. Gropper
March 1973

TASK

F DEVELOP DIAGNOSTIC AND EVALUATIVE TESTS

CONTENTS

page

F

STEPS

F.1 Plan the development of tests 1

SUB-STEPS

F.1.1 Determine the size of the instructional unit for which a set of tests will be developed 5

F.1.2 Make decisions about the types of tests (and their number) in each set of tests which will be developed for each instructional unit 11

F.1.3 Determine whether actual tests or storyboards for tests will be developed for each instructional unit 19

F.2 Develop tests for proficiency at criterion behavior which is taught in each instructional unit 27

F.2.1 Determine whether to test for product and/or for process 31

F.2.2 Decide on the sample of criterion behavior to be covered in a criterion test 41

F.2.3 Develop criterion test items or procedures 49

F.2.4 Develop subcriterion test items or procedures 71

F.3 Develop diagnostic tests to identify reason for failure to acquire the criterion behavior taught in each instructional unit — 79

F.3.1 Plan the use of results on criterion test items and on subcriterion test items in diagnosing type of learning failure that has occurred — 83

F.3.2 Develop diagnostic interview procedures to provide supplementary diagnostic information about results on criterion test items and subcriterion test items — 109

F.3.3 Develop diagnostic test items designed to identify specific types of learning failures — 131

F.4 Try out and revise testing procedures — 147

F.4.1 Administer testing procedures to a sample of the target audience as a means of assessing and upgrading the mechanics of the testing procedures — 151

F.4.2 Perform an analysis of the target audience by administering the revised tests — 161

STEP F.1

F.1 Plan the development of tests.

F.1.1 Determine the size of the instructional unit for which a set of tests will be developed.

F.1.2 Make decisions about the types of tests (and their number) in each set of tests which will be developed for each instructional unit.

F.1.3 Determine whether actual tests or storyboards for tests will be developed for each instructional unit.

STEP F.1

OVERVIEW

INPUT	ACTION	OUTPUT
Statement of objectives + simulation decisions	Plan the number and types of tests to be developed	DECISIONS about tests to be developed

	INPUT	ACTION	OUTPUT
F.1.1	Statements of criterion objectives on FORM D.2(1), grouped by lessons i	Determine the number of lessons to consider in making plans for development of tests ii	DECISION: -LESS than a lesson, -A lesson, or -MORE than a lesson as a testing unit iii
F.1.2	Statement of subcriterion objectives on FORM D.2(1), grouped by criterion objectives iv	Determine the number of subcriterion tests needed per set of tests devoted to an instructional unit v	DECISION: As many subcriterion tests per criterion objective (the instructional unit) as there are subcriterion objectives vi
F.1.3	Criterion behavior in each identified instructional unit + simulation decision (Step E) for it vii	Assess cost of preparing and trying out test of criterion behavior using required media; decide whether costs are acceptable viii	DECISION to develop: -Actual tests OR -Storyboard for tests ix

STEP F.1

PAGE INDEX

	CRITERIA FOR IDENTIFYING INPUTS	ACTION TO BE TAKEN	STANDARD FOR OUTPUTS	FORMS TO USE
F.1.1	-MATRIX: Unit of instructional material to be tested 8	-MATRIX: How many tests to develop per lesson 8		SUMMARY OF PROCEDURES . . . 9
F.1.2	-MATRIX: Three types of test 14	-MATRIX: How many types of test to prepare 15		SUMMARY OF PROCEDURES . . . 17
F.1.3		-MATRIX: When to develop actual tests vs. storyboard . . . 22		SUMMARY OF PROCEDURES . . . 23

PREVIEW OF THE NEXT SubSTEP

YOUR PRODUCT	*Plans to develop testing units for each criterion behavior (resulting in lesson units having as many testing units as criterion behaviors covered in them).*
WHAT YOU WILL WORK FROM	(1) Statements of criterion objectives grouped by lesson units.
WHAT YOU WILL DO	(1) Determine the number of criterion objectives covered in a lesson unit. (2) Decide whether the testing unit will be less than a lesson unit, a lesson unit, or two or more lesson units.
FORMS YOU WILL USE	None

DESCRIPTION OF Sub-STEP	F.1.1

INPUT	ACTION	OUTPUT
Statements of criterion objectives on FORM D.2(1) grouped by lessons i	Determine the number of lessons to consider in making plans for development of tests ii	DECISION: -LESS than a lesson, -A lesson, or -MORE than a lesson as a TESTING unit iii

Job Aid Contents

CRITERIA FOR IDENTIFYING INPUTS	ACTION TO BE TAKEN	STANDARD FOR OUTPUTS	FORMS TO USE
-MATRIX: Unit of instructional material to be tested 8	-MATRIX: How many tests to develop per lesson 8		SUMMARY OF PROCEDURES . . . 9

Required Materials

COMPLETED MATERIALS	STEP	COMPLETED FORMS	STEP	BLANK FORMS
		FORM D.2(1) carried forward from	E.1 E.2	

Sub-STEP F.1.1

JOB DIAGRAM

INPUT	ACTION	OUTPUT
Statements of criterion objectives on FORM D.2(1) grouped by lessons — i	Determine the number of lessons to consider in making plans for development of tests — ii	DECISION: -LESS than a lesson, -A lesson, or -MORE than a lesson as a TESTING unit — iii

TESTING PLANS

A single lesson covers more than one criterion objective — i.a	Treat instruction (in a lesson) covering each criterion objective as a unit for testing — ii.a	As many tests per lesson as there are criterion objectives — iii.a
A single lesson covers only one criterion objective — i.b	Treat the entire lesson as the unit for testing — ii.b	Only one test per lesson — iii.b
Several lessons cover only one criterion objective — i.c	Treat all the lessons required to cover a criterion objective as the unit for testing — ii.c	Only one test for all the lessons required to cover the criterion objective — iii.c

JOB PROCEDURES

	page
Size of instructional unit to use as a basis for test development	8
How many tests to develop per lesson	8
SUMMARY OF JOB PROCEDURES	9

F.1.1

IDENTIFICATION MATRIX

CRITERIA FOR IDENTIFYING APPROPRIATE INSTRUCTIONAL UNIT SIZE AS A BASIS FOR TEST DEVELOPMENT

APPROPRIATENESS OF UNIT	APPROPRIATE Instructional Unit to be tested	INAPPROPRIATE Instructional Unit to be tested
CRITERIA	*All the instructional material related to each CRITERION objective can range from:*	*Instructional material related to less than or part of a criterion objective*
	-less than a single lesson, to -multiple lessons	

F.1.1

DECISION MATRIX

DETERMINING HOW MANY TESTS TO DEVELOP PER LESSON

CONDITIONS	A single lesson covers more than one criterion objective	A single lesson covers only one criterion objective	It takes several lessons to cover a single criterion objective
ACTION TO TAKE	*Treat instruction covering each criterion objective as the unit to be tested* *Develop as many tests per lesson as there are criterion objectives*	*Treat instruction covering the one criterion objective as the unit to be tested* *Develop only one set of tests for that lesson*	*Treat instruction covering only complete criterion objective as the unit to be tested* *Develop only one set of tests for all the lessons it takes to cover the criterion objective*

F.1.1

ILLUSTRATION SUMMARIZING PROCEDURES FOR DETERMINING HOW MANY INSTRUCTIONAL UNITS THERE ARE AND HOW MANY TESTS WILL BE NEEDED

#1

a. Inspect the collection of "objectives" forms ~~[FORM D.2(1)] grouped~~ by lessons

b. Identify the number of criterion objectives covered by each lesson:

-None
-Only one
-More than one

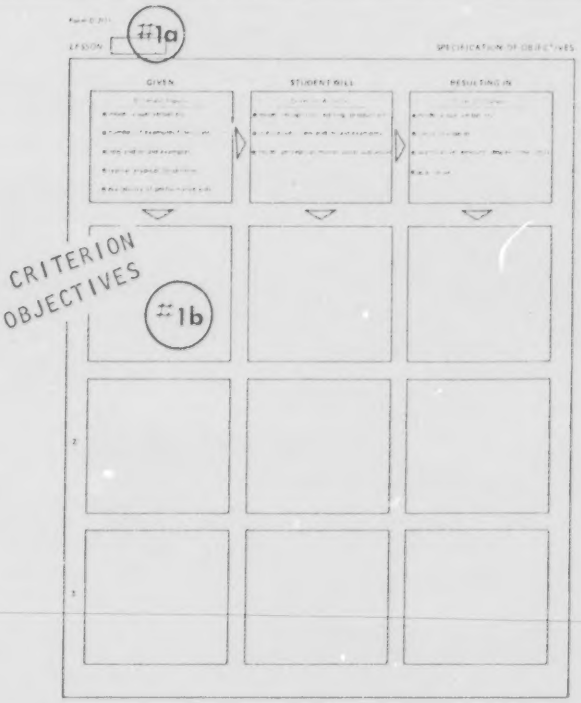
#2

a. Make plans to develop a set of tests for each criterion objective identified

b. Make plans to have as many sets of tests devoted to a lesson as there are criterion objectives:

-None
-Only one
-More than one

FORM D.2(1)



PREVIEW OF THE NEXT SubSTEP

YOUR PRODUCT	*Plans to develop a sub-criterion test for each sub-criterion objective.*
WHAT YOU WILL WORK FROM	(1) Statements of sub-criterion objectives; all those relevant to a given criterion behavior grouped together.
WHAT YOU WILL DO	(1) Determine how many sub-criterion tests will be prepared for each instructional unit.
FORMS YOU WILL USE	None

DESCRIPTION OF Sub-STEP	F.1.2

INPUT	ACTION	OUTPUT
Statement of sub-criterion objectives on FORM D.2(1) grouped by criterion objectives iv	Determine the number of sub-criterion tests needed per set of tests devoted to an instructional unit v	DECISION: As many sub-criterion tests per criterion objective (the instructional unit) as there are sub-criterion objectives vi

Job Aid Contents

CRITERIA FOR IDENTIFYING INPUTS	ACTION TO BE TAKEN	STANDARD FOR OUTPUTS	FORMS TO USE
-MATRIX: Three types of test 14	-MATRIX: How many types of test to prepare 15		SUMMARY OF PROCEDURES . . 17

Required Materials

COMPLETED MATERIALS	STEP	COMPLETED FORMS	STEP	BLANK FORMS
Identification of instructional unit	F.1.1	FORM D.2(1) carried forward from	F.1.1	

Sub-STEP F.1.2

JOB DIAGRAM

INPUT	ACTION	OUTPUT
Statement of sub-criterion objectives on FORM D.2(1 grouped by criterion objectives iv	Determine the number of sub-criterion tests needed per set of tests devoted to an instructional unit v	DECISION: As many sub-criterion tests per criterion objective (the instructional unit) as there are sub-criterion objectives vi

on FORM D.2(1)		DECISION
Only a criterion objective recorded iv.a	Plan on development only of a criterion test v.a	A criterion test to be developed vi.a
Associated with a criterion objective: -Subcriterion objective(s) recorded -No preparatory objective(s) recorded iv.b	In addition to above: Plan on development of a subcriterion test to cover subcriterion objectives v.b	-A criterion test to be developed + -As many subcriterion tests as there are subcriterion objectives vi.b
Associated with a criterion objective: -Subcriterion objective(s) recorded -Preparatory objective(s) recorded iv.c	In addition to above: Plan on development of a preparatory test to cover preparatory objectives v.c	-A criterion test to be developed + -As many subcriterion tests as there are subcriterion objectives + -As many preparatory tests as there are preparatory objectives vi.c

JOB PROCEDURES

	page
Types of tests to develop for each instructional unit	14
How many "preparatory" tests to develop per set of tests for each instructional unit	15
SUMMARY OF PROCEDURES	17

F.1.2

IDENTIFICATION MATRIX

CRITERIA FOR IDENTIFYING THE TYPES OF TESTS TO BE DEVELOPED FOR EACH INSTRUCTIONAL UNIT

TYPES OF TESTS	CRITERION TEST	SUBCRITERION TEST	PREPARATORY TEST
CRITERIA	*A test relevant to the criterion objective covered in the instructional unit*	*A test covering each and every subcriterion objective associated with the criterion objective:*	*A test covering all preparatory objectives (testing for component skills)*

EXAMPLES

Criterion Behavior is to compute the size of the buoyant force exerted on a submerged object (by submerging the object and weighing the overflow)	*The criterion test would require the student to exhibit the criterion behavior (as stated to the left)*	*A subcriterion test might call for the student to state the relationship between the weight of the overflow and the magnitude of the buoyant force (i.e., produce a formula)*	*A test of component skills might call for the student (given an arrow representing an upward force and an arrow representing a downward force) to select the one representing the "buoyant" force*

F.1.2

DECISION MATRIX

DETERMINING HOW MANY PREPARATORY TESTS TO PREPARE PER SET OF TESTS DEVOTED TO EACH INSTRUCTIONAL UNIT

CONDITIONS	-NO subcriterion or preparatory objectives have been prepared relating to a criterion objective	-Subcriterion objectives have been prepared relating to a criterion objective AND -NO preparatory objectives have been prepared relating to a criterion objective (or to a subcriterion objective)	-Preparatory objectives (covering component skills) have been prepared relating to: ••A criterion objective, or to a ••Subcriterion objectives relating to the criterion objective
ACTION TO TAKE	*-Plan to include in the set of tests:* *••Only a criterion test*	*IN ADDITION TO: A CRITERION TEST* *-Plan to include in the set of tests:* *••Only a subcriterion test* *-Plan as many such tests as there are subcriterion objec*	*IN ADDITION TO: A CRITERION TEST AND SUBCRITERION TESTS* *-Plan to include in the set of tests:* *••Preparatory tests* *-Plan as many such tests as there are preparatory objectives*

F.1.2

ILLUSTRATION SUMMARIZING PROCEDURES INVOLVED IN DETERMINING HOW MANY OF EACH TYPE OF TEST TO INCLUDE IN EACH SET OF TESTS PER INSTRUCTIONAL UNIT

#1

a. Inspect the collection of "objectives" forms [FORM D.2(1)] associated with each criterion objective

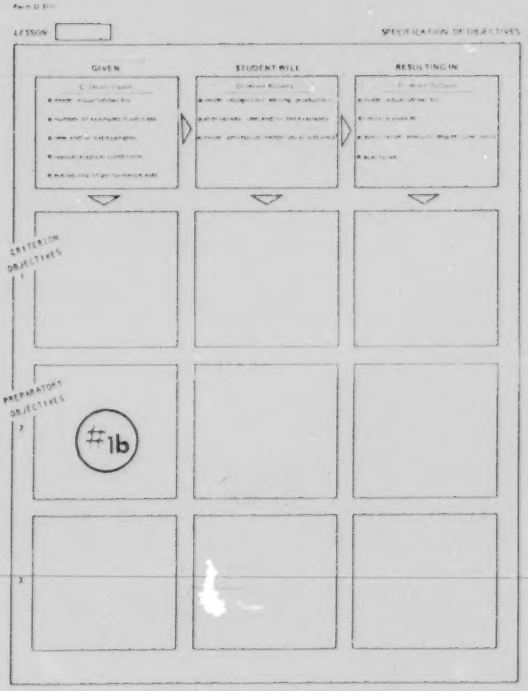
b. Identify the number of sub-objectives associated with each criterion objective

#2

a. Make plans to develop a test for each sub-objective

••Subcriterion objective
••Preparatory objective

FORM D.2(1)



PREVIEW OF THE NEXT SubSTEP

YOUR PRODUCT	*A decision to develop tests using media required by criterion behavior or to use storyboard tests because they are cheaper.*
WHAT YOU WILL WORK FROM	(1) Classification of the criterion behavior to be practiced in an instructional unit. (2) Decisions which specify that a simulation of the criterion behavior will be practiced.
WHAT YOU WILL DO	(1) Assess the costs of developing and using tests using the media needed for testing criterion behavior or a simulated version of it.
FORMS YOU WILL USE	None

DESCRIPTION OF Sub-STEP	F.1.3

INPUT	ACTION	OUTPUT
Criterion behavior in each identified instructional unit + simulation decisions (Step E) for it vii	Assess cost of preparing and trying out test of criterion behavior using required media; decide whether costs are acceptable viii	DECISION to develop: -Actual tests OR -Storyboard for tests ix

Job Aid Contents

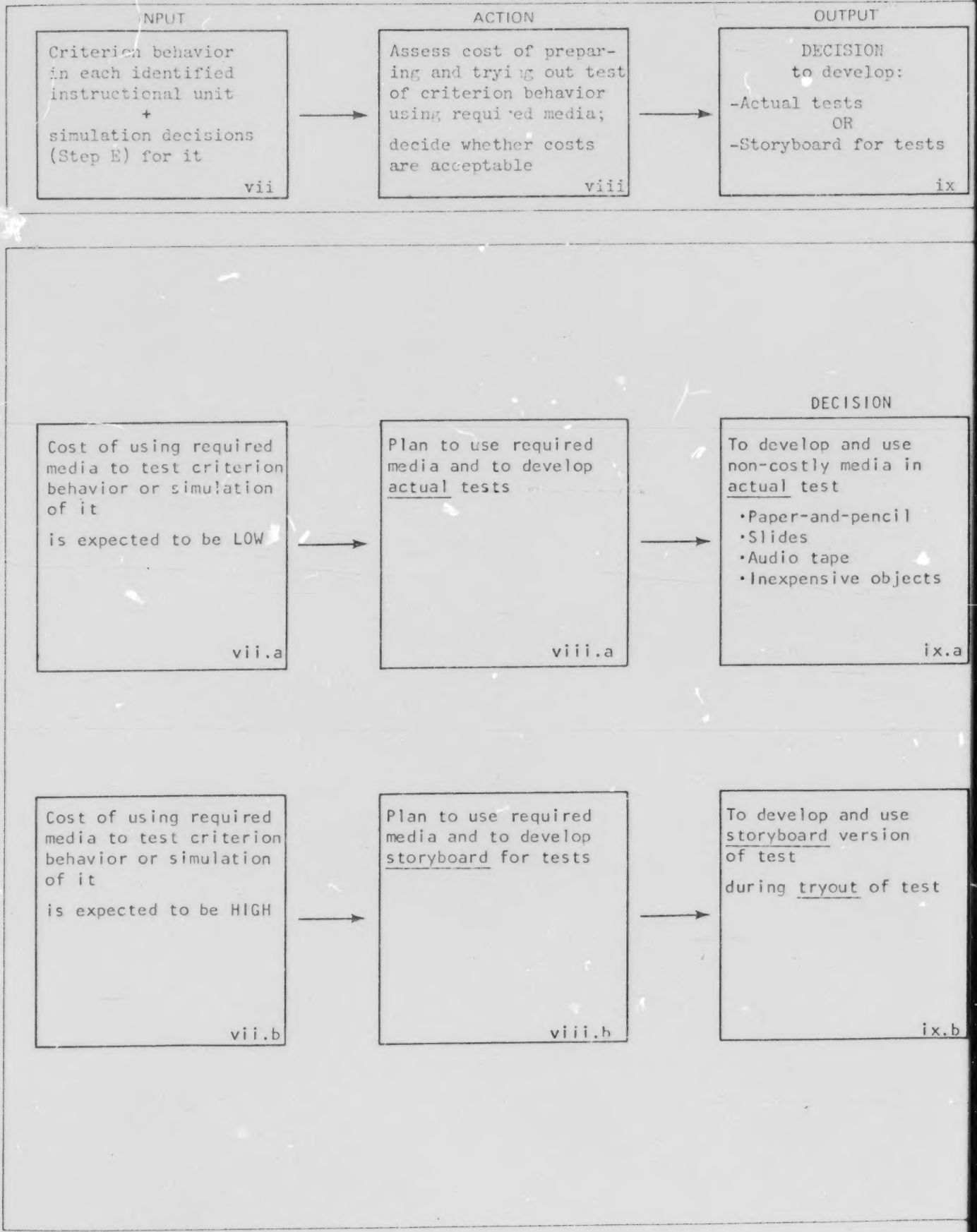
CRITERIA FOR IDENTIFYING INPUTS	ACTION TO BE TAKEN	STANDARD FOR OUTPUTS	FORMS TO USE
	-MATRIX: When to develop actual tests vs. storyboard . . . 22		SUMMARY OF PROCEDURES . . . 23

Required Materials

COMPLETED MATERIALS	STEP	COMPLETED FORMS	STEP	BLANK FORMS
Identification of instructional unit	F.1.1	FORM E.1(1) Simulation decision	E.2.3	
Identification of types of tests to use for unit	F.1.2			

Sub- STEP F.1.3

JOB DIAGRAM


INPUT
ACTION
OUTPUT
Criterion behavior in each identified instructional unit
+
simulation decisions (Step E) for it
vii
Assess cost of preparing and trying out test of criterion behavior using required media;
decide whether costs are acceptable
viii
DECISION to develop:
-Actual tests
OR
-Storyboard for tests
ix
DECISION
Cost of using required media to test criterion behavior or simulation of it
is expected to be LOW
vii.a
Plan to use required media and to develop actual tests
viii.a
To develop and use non-costly media in actual test
•Paper-and-pencil
•Slides
•Audio tape
•Inexpensive objects
ix.a
Cost of using required media to test criterion behavior or simulation of it
is expected to be HIGH
vii.b
Plan to use required media and to develop storyboard for tests
viii.b
To develop and use storyboard version of test
during tryout of test
ix.b

JOB PROCEDURES

	page
Determining whether to develop actual tests or storyboards for tests	22
SUMMARY OF PROCEDURES	23

F.1.3

DECISION MATRIX

DETERMINING WHETHER TO DEVELOP EITHER ACTUAL TESTS AND/OR STORYBOARDS FOR TESTS

CONDITIONS	-Preparing a test of criterion behavior OR -Preparing a test of simulated criterion behavior results in LOW costs due to combinations of: ••Time ••Money ••Facilities ••Personnel ••Materials	-Preparing a test of criterion behavior OR -Preparing a test of simulated criterion behavior results in HIGH costs due to combinations of: ••Time ••Money ••Facilities ••Personnel ••Materials
ACTION TO TAKE	*Plan to develop an ACTUAL TEST*	*Plan to develop test in STORYBOARD FORM*≠*

**Following tryout of the adequacy of a test in storyboard form, the actual test would be developed.*

≠Subsequent sections do not provide details of procedures for developing and trying out materials in storyboard form.

EXAMPLES	-Use of relatively inexpensive media (paper and pencil, slides, audio tape, inexpensive objects, etc.) e.g., *the criterion behavior is "the identification of types of leaves." Slides of the leaves to be identified are prepared, since it is not costly to do so, nor is there likely to be a need to revise the "shots" of the leaves*	-Use of costly media (film, TV, complex simulators, etc.) e.g., *criterion behavior for a developer of instructional materials involves "preparation of training films"; a test of this behavior using actual film could be costly; therefore, the film test is first scripted (and tried out in that form) before it is actually produced*

F.1.3

ILLUSTRATION SUMMARIZING PROCEDURES FOR DETERMINING WHETHER, FOR EACH INSTRUCTIONAL UNIT, TO DEVELOP ACTUAL TESTS OR STORYBOARD FOR TESTS

#1

FOR EACH CRITERION OBJECTIVE (which defines the instructional unit)

a. Inspect the collection of "objectives" forms [FORM D.2(1)] associated with that criterion objective

b. Inspect the simulation decision [on FORM E.1(1)] made with respect to the criterion behavior covered in the statement of the objective

#2

a. Assess the costs of using the media required for an actual test of the criterion behavior (or the simulated version decided on)

b. Plan the use of actual tests or storyboard version during tryout of test based on cost considerations

STEP F.1

COMPLETION CHECKLIST

	IDENTIFIED	PERFORMED	PRODUCED	FORMS COMPLETED
F.1.1	The number of instructional units for which tests will be needed			
F.1.2	The number of different types of tests that will be necessary for instructional unit			
F.1.3		Decided whether to develop actual tests or a storyboard for tests		

STEP F.2

F.2 Develop tests for proficiency at criterion behavior which is taught in each instructional unit.

F.2.1 Determine whether to test for product and/or for process.

F.2.2 Decide on the sample of criterion behavior to be covered in a criterion test.

F.2.3 Develop criterion test items or procedures.

F.2.4 Develop subcriterion test items or procedures.

STEP F.2

OVERVIEW

INPUT	ACTION	OUTPUT
All analysis results for each criterion behavior + statement of objectives + simulation decisions + testing decisions	Develop tests of proficiency for each criterion behavior	Criterion tests + Subcriterion tests assessing proficiency in criterion behavior and associated Subcriterion behaviors

F.2.1	Grouped statement of objectives + associated learning analysis i	Decide whether to test for criterion or subcriterion outputs ii	DECISION to test for -Criterion output (product) -Subcriterion output (process) iii
F.2.2	Statements of objectives + task analysis forms for each criterion behavior iv	Determine how much of criterion behavior to sample on test v	DECISION: -Number of INPUT classes to sample -Number of test items per class to develop vi
F.2.3	Statement of criterion objective + sampling decision + simulation decision vii	Prepare criterion test item(s) viii	CRITERION TEST item(s) containing: -GIVEN -STUDENT WILL -RESULTING IN ix
F.2.4	Statement of subcriterion objective + sampling decision + simulation decision x	Prepare subcriterion test item(s) xi	Subcriterion test items containing: -GIVEN -STUDENT WILL -RESULTING IN xii

STEP F.2

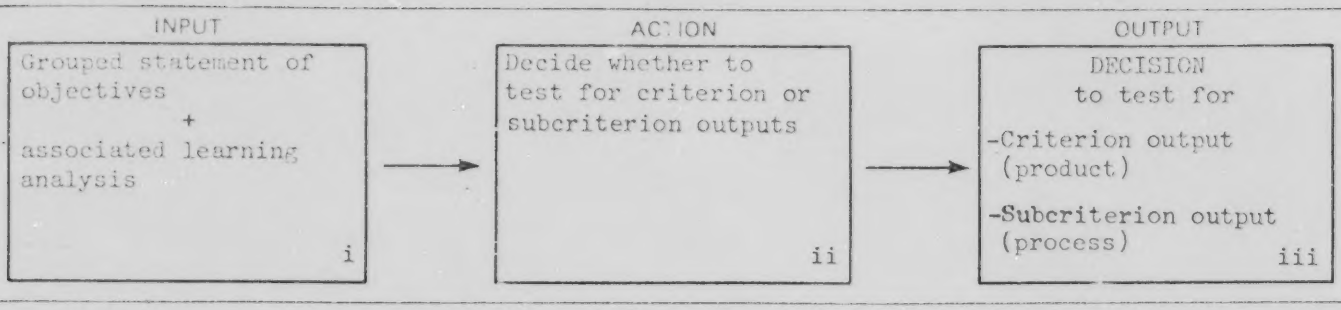
PAGE INDEX

	CRITERIA FOR IDENTIFYING INPUTS	ACTION TO BE TAKEN	STANDARD FOR OUTPUTS	FORMS TO USE
F.2.1	-MATRIX: What is "product" and "process" . . . . 34 -MATRIX: Three types of outputs to be tested . . . . . 35 -MATRIX: Functions served by tests 38	-MATRIX: When to test for criterion and for preparatory outputs 38		SUMMARY OF PROCEDURES . . 39
F.2.2	-MATRIX: Two required sampling decisions . . . 44	-MATRIX: How many INPUT classes to sample . . . . 45 -MATRIX: How many test items per class to develop 46		SUMMARY OF PROCEDURES . . 47
F.2.3	-MATRIX: Types of test items . . 51-57	-MATRIX: Making scoring "objective" . 59-61 -MATRIX: What to include in items and where to get information . . 67	-MATRIX: Adequacy of test items 69	FORM F.2(1) . . 65 SUMMARY OF PROCEDURES . . 68
F.2.4	-MATRIX: Difference between criterion and subcriterion test item 74			FORM F.2(1) SUMMARY OF PROCEDURES . . 75

PREVIEW OF THE NEXT SubSTEP

YOUR PRODUCT	*Plans to develop tests which assess proficiency at criterion behavior and, when necessary, proficiency at intermediate goals.*
WHAT YOU WILL WORK FROM	(1) Grouped statements of objectives. (2) Grouped analyses of criterion behavior.
WHAT YOU WILL DO	(1) Identify anticipated length of instructional units (2) Based on anticipated lesson size, determine whether to prepare tests on criterion behavior and/or on criterion progress.
FORMS YOU WILL USE	None

DESCRIPTION OF Sub-STEP	F.2.1



Job Aid Contents

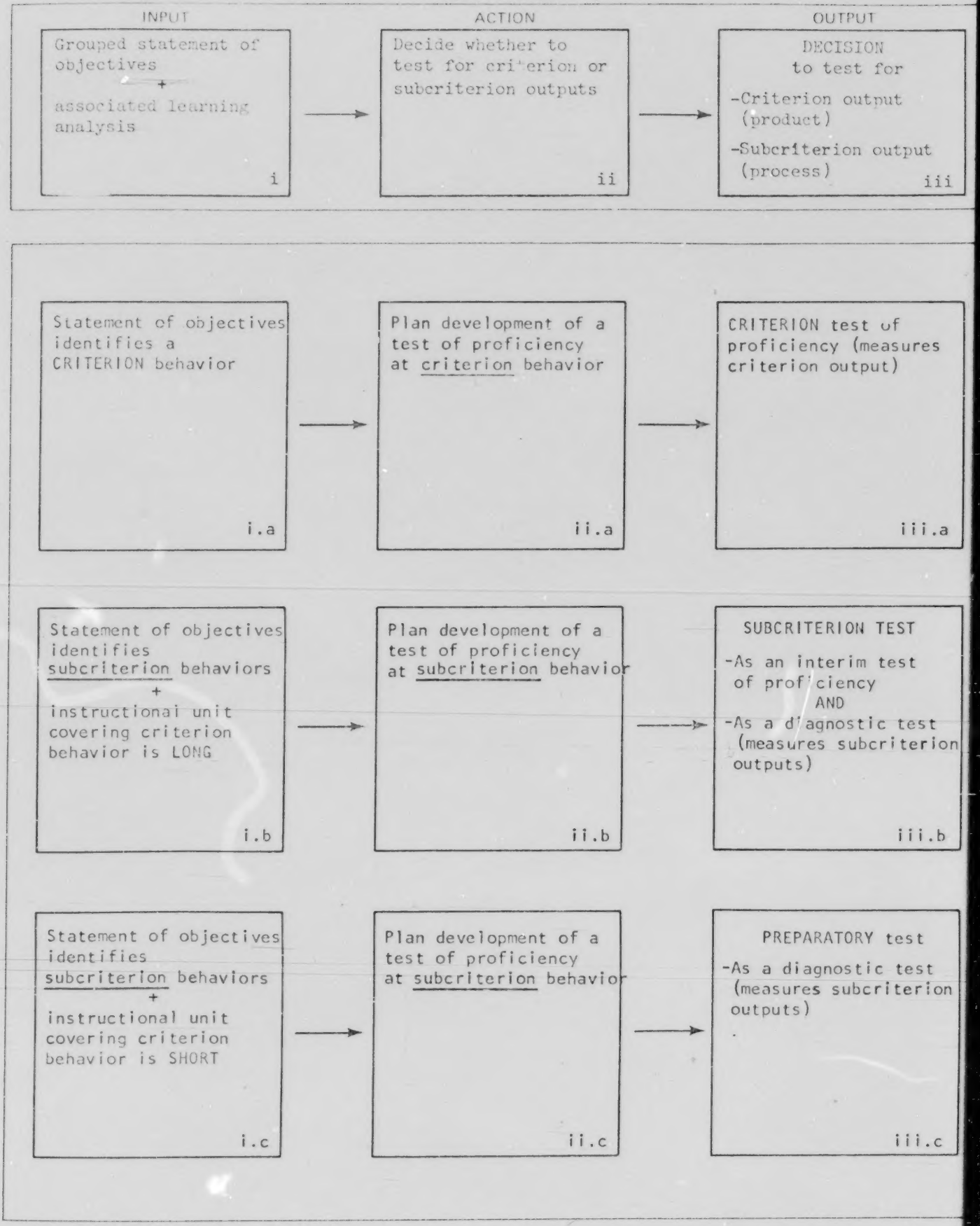
CRITERIA FOR IDENTIFYING INPUTS	ACTION TO BE TAKEN	STANDARD FOR OUTPUTS	FORMS TO USE
-MATRIX: What is "product" and "process" . . . . 34 -MATRIX: Three types of outputs to be tested . . . . . 35 -MATRIX: Functions served by tests 38	-MATRIX: When to test for criterion and for subcriterion outputs 38		SUMMARY OF PROCEDURES . . . 39

Required Materials

COMPLETED MATERIALS	STEP	COMPLETED FORMS	STEP	BLANK FORMS
Test planning decisions	F.1.3	Collection of forms carried forward from	F.1.3	

Sub- STEP F.2.1

JOB DIAGRAM



BACKGROUND INFORMATION

	page
Distinction between "product" and "process"	34
Three types of outputs	35

F.2.1

CRITERIA FOR DISTINGUISHING BETWEEN "PRODUCT" AND "PROCESS"

IDENTIFICATION MATRIX

TERM	PRODUCT	PROCESS
CRITERIA	*A product is defined as the final OUTPUT of a criterion behavior*	*Process is defined by and measured by the OUTPUTS or products of unmodified preparatory behaviors which make up the criterion behavior*
IMPLICATIONS	The testing of behavior is therefore always made in terms of outputs or products; but outputs may be at differing stages of completion. In effect, testing for process means to test products which are partial or preliminary	
EXAMPLES	MATHEMATICS: *adding two-digit numbers*	
	e.g., *a "sum" is the final output of adding numbers and can be used to test for the criterion behavior*	e.g., *the answer in the digits (only) column would be the output of the preparatory behavior and can be used to test for the process of "adding"*
	ENGLISH: *writing an essay*	
	e.g., *a completed essay is the final output or product and can be used to assess the criterion behavior of writing an essay*	e.g., *an outline could be used as the output of a preparatory behavior and used to assess or test the process of "writing an essay"*

F.2.1

IDENTIFICATION MATRIX

CRITERIA FOR IDENTIFYING THREE TYPES OF OUTPUTS--FINAL OR PREPARATORY

TYPES OF OUTPUTS	I	II	III
CRITERIA	*Outputs are:* *-Tangible* *-Observable* *-Permanent* *·Objects, things* *·Written words or symbols*	*Outputs are:* *-Non-tangible* *-Observable* *-Transient* *·Motor behavior*	*Outputs are:* *-Cognitive* *-Unobservable* *-Transient* *·Identifications* *·Thoughts* *·Reading*
IMPLICATIONS FOR TESTING	This type of output can be measured: -Directly -At any time	To be measured, this type of output: -Has to be measured at the time it occurs -Has to be recorded for subsequent measurement	This type of output has to be measured: -Indirectly -By requiring the student to do something else
EXAMPLES	*e.g., the product in multiplication problems* *e.g., a chemical distillate (output of an experiment)* *e.g., an essay* *e.g., a painting*	*e.g., speech (can be recorded on tape as a permanent record)* *e.g., athletic outcomes (movements as outputs or results of movements as outputs)* *·A dive* *·Height of a pole vault (can be recorded on film)* *e.g., using instruments in sciences* *·How held or moved*	*e.g., reading (to be tested, questions have to be asked about what was read)* *e.g., identifications (to be tested, the student has to be required to point to something or to label something)*

JOB PROCEDURES

	page
Functions served by testing	38
Determining types of tests to develop	38
SUMMARY OF PROCEDURES	39

F.2.1

IDENTIFICATION MATRIX

CRITERIA IDENTIFYING THE FUNCTIONS SERVED BY TESTING: FOR CRITERION OUTPUTS AND FOR SUBCRITERION OUTPUTS

TYPE OF TEST	Test of CRITERION outputs	Test of SUBCRITERION outputs
CRITERIA	*-Test of student proficiency at criterion behavior*	*-Test of student proficiency at subcriterion behaviors*
FUNCTIONS SERVED	Assess the adequacy of the <u>total</u> instructional unit to produce a predetermined level of student proficiency in criterion behavior	Assess the adequacy of a portion of the instructional unit to produce a predetermined level of student proficiency in a subcriterion behavior: ••Can be used to measure interim progress; AND/OR ••Can be used diagnostically to record weaknesses in parts of instructional methods

F.2.1

DECISION MATRIX

DETERMINING WHEN TO TEST FOR: CRITERION AND SUBCRITERION OUTPUTS

CONDITIONS	For <u>ANY</u> criterion output	-When instructional materials for a given criterion behavior are spread over <u>several</u> lessons AND/OR -When it is <u>difficult</u> to acquire the criterion behavior
ACTION TO TAKE	*Plan to develop an instrument to test for proficiency of CRITERION behavior*	*Plan to develop an instrument to test for:* *--Interim progress; or for* *--Diagnostic evidence of learning difficulties*

F.2.1

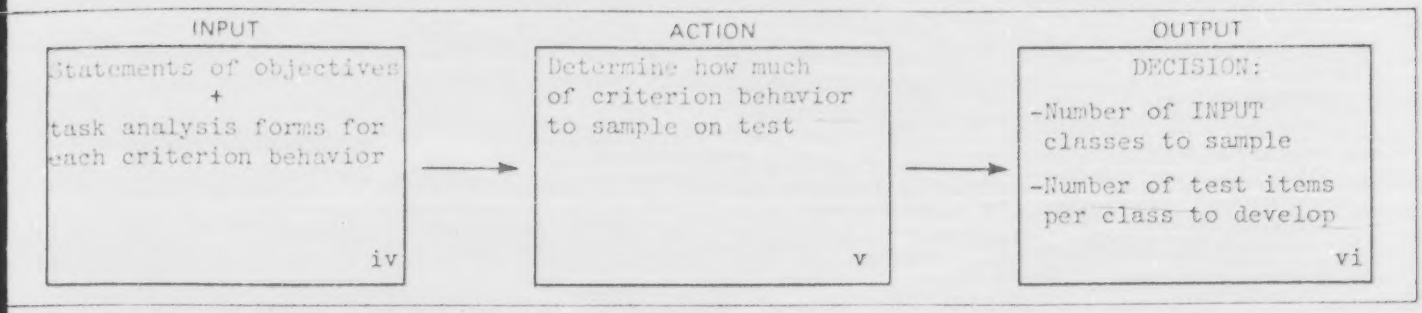
ILLUSTRATION SUMMARIZING PROCEDURES INVOLVED IN DECIDING WHEN TO TEST FOR "PRODUCT" (CRITERION OUTPUT) AND FOR "PROCESS" (PREPARATORY OUTPUT)

#1	#2
a. Inspect the decision about size of instructional unit made in Sub-STEP F.1.1 b. Inspect the collection of analysis forms and statements of objectives for determining anticipated number of lessons likely to be required to teach a given criterion behavior.	a. Plan to develop a criterion test for proficiency at the criterion behavior b.1. For LONG instructional units, plan to use preparatory tests (associated with unmodified preparatory objectives) as: ··Interim tests of proficiency ··Diagnostic tests b.2. For SHORT instructional units, plan to use preparatory tests (associated with unmodified preparatory objectives) as: ··Diagnostic tests (only)

PREVIEW OF THE NEXT SubSTEP

YOUR PRODUCT	*A decision as to how many of the INPUT classes or ACTION classes to sample on a test as well as how many test items for each class to develop.*
WHAT YOU WILL WORK FROM	(1) Statements of objectives (2) Task analyses for each criterion behavior.
WHAT YOU WILL DO	(1) Determine how much of the criterion behavior to test for in a test.
FORMS YOU WILL USE	None

DESCRIPTION OF Sub-STEP	F.2.2



Job Aid Contents

CRITERIA FOR IDENTIFYING INPUTS	ACTION TO BE TAKEN	STANDARD FOR OUTPUTS	FORMS TO USE
-MATRIX: Two required sampling decisions . . . 44	-MATRIX: How many INPUT classes to sample . . . . . 45 -MATRIX: How many test items per class to develop 46		SUMMARY OF PROCEDURES . . . 47

Required Materials

COMPLETED MATERIALS	STEP	COMPLETED FORMS	STEP	BLANK FORMS
Decision to test for product/process	F.2.1	Collected forms for criterion behavior carried forward from	F.2.1	

Sub-STEP F.2.2

JOB DIAGRAM

INPUT	ACTION	OUTPUT
Statement of objectives + task analysis forms for each criterion behavior iv	Determine how much of criterion behavior to sample on test v	DECISION: -Number of INPUT classes to sample -Number of test items per class to develop vi

STATEMENT OF OBJECTIVES identifies

		PLANS
-Relatively more complex performance: ••There are many classes, and/or ••There are many members per class AND: -Performance involves TRANSFER (because of above conditions) iv.a	Plan to sample INPUT classes less heavily v.a	-All classes to be sampled BUT -Fewer members of classes of INPUTS to be sampled ••Fewer test items per class to be sampled vi.a
-Relatively less complex performance: ••There are few classes, and/or ••There are few members per class AND -Performance involves RECALL (because of above conditions) iv.b	Plan to sample INPUT classes more heavily v.b	-All classes to be sampled AND -More members of classes of INPUTS to be sampled ••More test items per class to be sampled vi.b

JOB PROCEDURES

	page
What and how much of criterion behavior to sample	44
Determining the sampling of classes of inputs	45
Determining how many test items to be developed per class of inputs to be sampled	46
SUMMARY OF PROCEDURES	47

*This section provides rules of thumb only. They are based on learning considerations primarily and "testing" considerations secondarily. Consult other works on "test development" concerning testing issues related to test length.

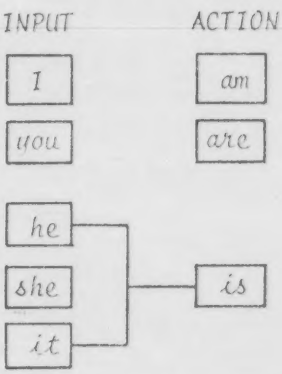
F.2.2

CRITERIA FOR IDENTIFYING TWO REQUIRED TYPES OF SAMPLING DECISIONS

IDENTIFICATION MATRIX

TYPES OF DECISIONS	WHAT elements of criterion behavior need to be sampled on a test?	HOW MANY test items should there be for each element decided on?
CRITERIA	*When they are identified in statements of objectives, should you sample:* *-All classes (INPUTS or ACTIONS) or just some of them* *-All members within each class, or just some of them*	*When it is decided how many classes and how many members within a class are to be sampled, should you:* *-Develop one or more test items per class* *-Develop one or more test items per member within a class*

ENGLISH - Pronoun/verb agreement



	On a criterion test should each of the three INPUT classes be represented; should all three members of the class of third person pronouns be represented?	On a criterion test how many test items should there be for each of the three classes; how many for each member within the class of third person pronouns?

F.2.2

DETERMINING WHEN TO SAMPLE ALL CLASSES OF INPUTS (THE "GIVEN" IN A STATEMENT OF OBJECTIVES)

DECISION MATRIX

CONDITIONS	The INPUTS described in the "given" portion of a statement of objectives involve (combinations of): -Relatively much content -Relatively difficult learning problems -Transfer AND The ACTIONS described in the "student will" portion of a statement of objectives involve (combinations of): -Relatively long chains -Relatively difficult learning problems -Transfer	The INPUTS described in the "given" portion of a statement of objectives involve (combinations of): -Relatively little content -Relatively easy learning problems -Recall AND The ACTIONS described in the "student will" portion of a statement of objectives involve (combinations of): -Relatively long chains -Relatively difficult learning problems -Transfer	The INPUTS described in the "given" portion of a statement of objectives involve (combinations of): -Relatively much content -Relatively difficult learning problems -Transfer AND The ACTIONS described in the "student will" portion of a statement of objectives involve (combinations of): -Relatively short chains -Relatively easy learning problems -Recall	The INPUTS described in the "given" portion of a statement of objectives involve (combinations of): -Relatively little content -Relatively easy learning problems -Recall AND The ACTIONS described in the "student will" portion of a statement of objectives involve (combinations of): -Relatively short chains -Relatively easy learning problems -Recall
ACTION TO TAKE	*-Plan to sample less than all members of each INPUT class* *-Plan to use transfer members of the INPUT class*	*-Plan to sample all members of the INPUT classes* *-Plan to require transfer members of the ACTION classes*	*-Plan to sample less than all members of each INPUT class* *-Plan to require all members of ACTION class*	*-Plan to sample all the members of INPUT and ACTION classes*
EXAMPLES	e.g., OBJECTIVE GIVEN: *A long prose paragraph* STUDENT WILL: *Analyze for style and characterize the style in his own words and rate it*	e.g., OBJECTIVE GIVEN: *The terms "restrictive" and "non-restrictive" clauses* STUDENT WILL: *Define them, give examples, and state the rule for punctuating the clauses*	e.g., OBJECTIVE GIVEN: *A long prose paragraph (not encountered in training) in part correctly and in part incorrectly punctuated* STUDENT WILL: *Edit the paragraph for punctuation*	e.g., OBJECTIVE GIVEN: *A personal pronoun* STUDENT WILL: *Provide the correct form of the present tense of the verb "to be"*

F.2.2

DETERMINING HOW MANY TEST ITEMS TO DEVELOP FOR EACH INPUT CLASS TO BE SAMPLED
(THE "GIVEN" IN A STATEMENT OF OBJECTIVES) (See previous page)

DECISION MATRIX

CONDITIONS	-A large number of classes of INPUTS to be sampled ··Three or more classes -Performance requirements are for: TRANSFER	-A large number of classes of INPUTS to be sampled ··Three or more classes -Performance requirements are for: RECALL	-A small number of classes of INPUTS to be sampled ··Two classes -Performance requirements are for: TRANSFER	-A small number of classes of INPUTS to be sampled ··Two classes -Performance requirements are for: RECALL
ACTION TO TAKE	*-Plan to develop:* *··Transfer items* *-Plan to develop:* *··The least number of such items, i.e., as few as one per each class*	*-Plan to develop:* *··Recall items* *-Plan to develop:* *··A minimum of one item per class* *··More items when the number of members within a class is large*	*-Plan to develop:* *··Transfer items* *··Recall items* *-Plan to develop multiple items for each class:* *··Two or more recall items* *··Two or more transfer items* *··More items when the number of members within a class is large*	*-Plan to develop:* *··Recall items* *-Plan to develop multiple items for each class:* *··Two or more items per class* *-When feasible (in terms of time requirements), have as many items as there are members of each class*

F.2.2

ILLUSTRATION SUMMARIZING PROCEDURES FOR DETERMINING THE SIZE OF THE SAMPLE OF CRITERION BEHAVIOR TO BE REPRESENTED IN THE CRITERION TEST

#1

FOR EACH CRITERION BEHAVIOR:

a. Inspect FORM D.2(1) (statement of objectives) for:

  a.1 Transfer/recall requirements

  a.2 Number of classes of inputs identified

  a.3 Number of members per class identified

b. Inspect the collection of task analysis forms [A.5(4) or (11)] for:

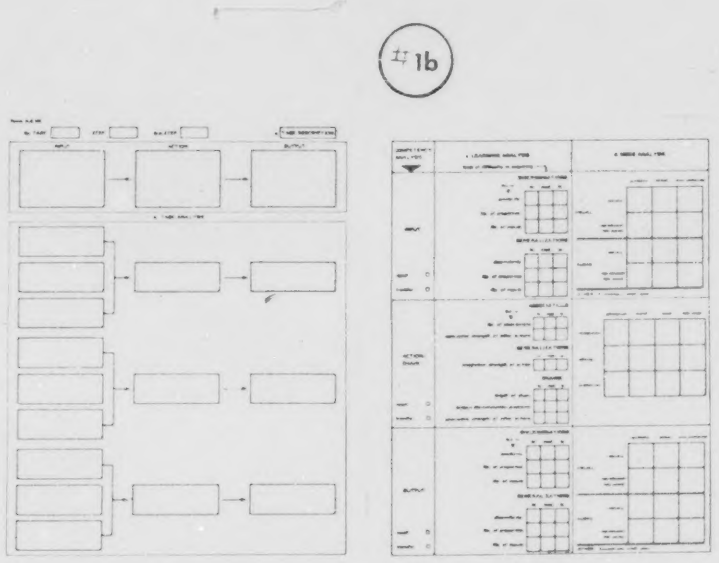
  -The length of chain involved

#2

a. Decide on the number of classes of INPUTS to be represented

b. Decide on the number of test items per class to be developed

FORM A.5(4) or A.5(11)

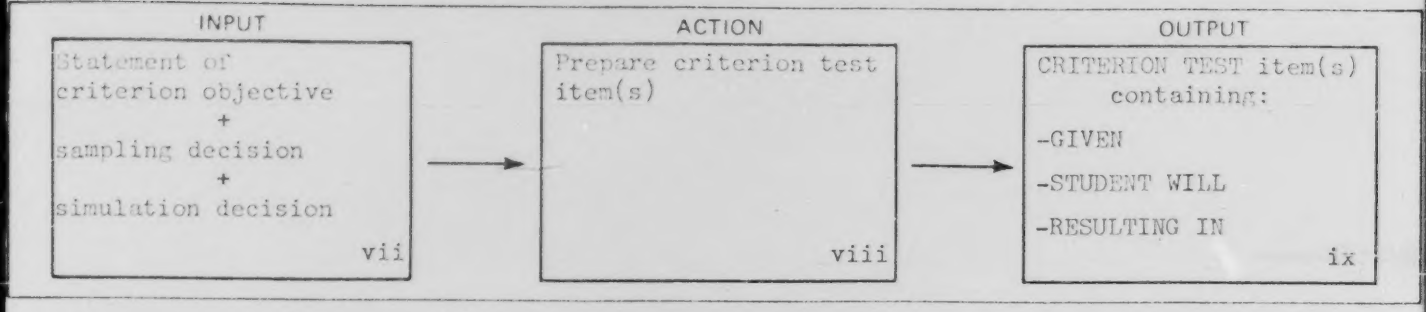


FORM D.2(1)

PREVIEW OF THE NEXT SubSTEP

YOUR PRODUCT	*The criterion test containing one or more items with specifications for an item spelling out: (1) what the student will be given; (2) what he will be expected to do; and (3) what he will be expected to produce.*
WHAT YOU WILL WORK FROM	(1) Statement of objectives re: each criterion behavior (2) Simulation decisions re: the objective (3) decision about how much behavior to sample and how many items to develop.
WHAT YOU WILL DO	(1) Prepare criterion test item(s).
FORMS YOU WILL USE	FORM F.2(1) for developing test items.

DESCRIPTION OF Sub-STEP	F.2.3



Job Aid Contents

CRITERIA FOR IDENTIFYING INPUTS	ACTION TO BE TAKEN	STANDARD FOR OUTPUTS	FORMS TO USE
-MATRIX: Types of test items . . 51-57	-MATRIX: Making scoring "objective" . 59-61 -MATRIX: What to include in items and where to get information . . . 67	-MATRIX: Adequacy of test items 69	FORM F.2(1) . . 65 SUMMARY OF PROCEDURES . . 68

Required Materials

COMPLETED MATERIALS	STEP	COMPLETED FORMS	STEP	BLANK FORMS
Sampling decision	F.2.2	Collection of FORMS carried forward from	F.2.2	FORM F.2(1)
Selection of simulation plan (where applicable)	E.2.3			

Sub-STEP F.2.3

JOB DIAGRAM

INPUT	ACTION	OUTPUT
Statement of criterion objective + sampling decision + simulation decision — vii	Prepare criterion test item(s) — viii	CRITERION TEST item(s) containing: -GIVEN -STUDENT WILL -RESULTING IN — ix

STATEMENT OF OBJECTIVES		CRITERION TEST ITEM ELEMENTS
GIVEN — vii.a	Translate into test item content — viii.a	INFORMATION TO BE GIVEN THE STUDENT — ix.a
STUDENT WILL — vii.b	Translate into test item content — viii.b	WHAT THE STUDENT IS EXPECTED TO DO — ix.b
RESULTING IN — vii.c	Translate into test item content — viii.c	-WHAT THE STUDENT IS EXPECTED TO TURN OUT -STANDARDS — ix.c

TYPES OF TEST ITEMS*

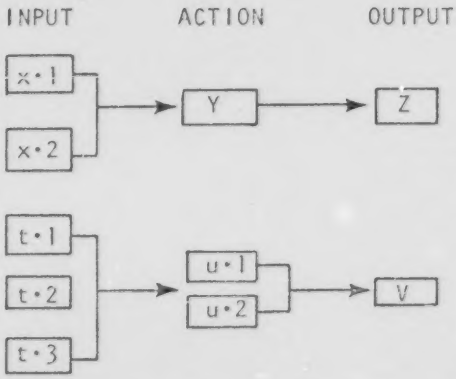
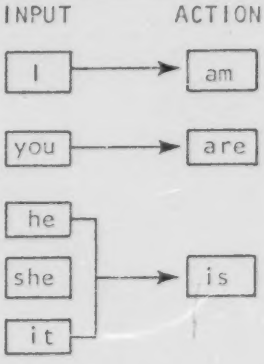
	page
How a test item samples criterion behavior	52
What a test item tests for	53
Test items requiring recall vs. transfer	54
Three elements to a test item	55
Three possible modes of response in test items	56
Desirable properties of test items	57

The phrase "test item" is meant to include any measurement situation (whether a paper-and-pencil test, an assignment involving objects and/or other people, a procedure) designed to assess acquisition or retention of criterion behavior.

F.2.3

IDENTIFICATION MATRIX

CRITERIA FOR IDENTIFYING HOW A TEST ITEM SAMPLES CRITERION BEHAVIOR

DIFFERENCE BETWEEN "TOTAL" CRITERION BEHAVIOR AND A TEST ITEM	"TOTAL" CRITERION BEHAVIOR represented in a task analysis diagram	A TEST ITEM testing criterion behavior
CRITERIA	*-A task analysis diagram represents all the classes (both of INPUTS and of ACTIONS) which the performer may encounter on a test or on the job* *-As such it does not reflect the real world in which only one member of a class is usually encountered*	*-A test item presents (usually) only one member of a class (INPUTS or ACTIONS) at a time* *-As such it reflects the real world in which criterion behavior requires dealing (usually) with one instance at a time*
GENERAL CASE		GIVEN: t•2 STUDENT WILL: Produce either u•1 or u•2
EXAMPLE		GIVEN: "he" STUDENT WILL: Use "is"

F.2.3

IDENTIFICATION MATRIX

	CRITERIA FOR IDENTIFYING WHAT A TEST ITEM TESTS FOR		
	1	2	3
WHAT IS TESTED FOR	Capability of exhibiting: the shortest possible chain, i.e., a simple association	Capability of exhibiting: a chain of intermediate length, i.e., a series of associations	Capability of exhibiting: a very long chain, i.e., a long series of associations
CRITERIA	*-The student is given one criterion INPUT in a test item* *-To be able to exhibit the chain, he must* *••Discriminate the input from other possible inputs* *••When the input belongs to a class of inputs, he must generalize to the test input (if not previously encountered)* *••Produce the action associated with the input (and not any other)* *or* *••Produce an alternative action (generalize) if another is allowable*	The intermediate case	*-The student is given one criterion INPUT in a test item* *-To be able to exhibit the chain, he must* *••Do everything in Column 1* *••Do everything in Column 1 repeatedly as each action produces an output which becomes the input for another action* *••Do all the repetitions in the correct sequence*
EXAMPLES	GIVEN: The INPUT - "the boy" STUDENT WILL: Produce the ACTION (saying) - "walks"	GIVEN: The INPUT - a quadratic equation and instructions to factor it STUDENT WILL: Take the series of actions involved in factoring	GIVEN: The INPUT - a specific novel to read and instructions to write an essay about it STUDENT WILL: Take the series of ACTIONS required to do this

F.2.3

IDENTIFICATION MATRIX

CRITERIA FOR IDENTIFYING TEST ITEMS OR TEST PROCEDURES INVOLVING RECALL VS. TRANSFER

TYPE OF REQUIREMENT	RECALL of INPUT and/or ACTION	TRANSFER for INPUT and/or ACTION
CRITERIA	*The "GIVEN" presents a student with an INPUT already encountered during instruction, i.e., an old input* *AND/OR* *The "STUDENT WILL" requires the student to recall an ACTION already practiced during instruction, i.e., an old action*	*The "GIVEN" presents a student with an INPUT not previously encountered during instruction, i.e., a new input* *AND/OR* *The "STUDENT WILL" requires the student to transfer to an ACTION not previously practiced during instruction, i.e., a new action*
EXAMPLES	ENGLISH - Noun/verb agreement	
	e.g., when he goes to work, the man *__ take a bus* *__ takes a bus* Both the INPUT -*man*- and the ACTION associated with it -*takes*- were used during instruction; therefore, both input RECALL and action RECALL are required	*e.g., when searching for food, the tiger* *__ hunts at night* *__ hunt at night* Both the INPUT -*tiger*- and the ACTION associated with it -*hunts*- were NOT used during instruction; therefore, both input TRANSFER and action TRANSFER are required

F.2.3
IDENTIFICATION MATRIX

CRITERIA FOR IDENTIFYING THREE CONTENT ELEMENTS FOR A TEST ITEM OR TEST PROCEDURE: (THEY PARALLEL THE THREE ELEMENTS IN A STATEMENT OF OBJECTIVES):

TEST ELEMENTS	"GIVEN"	"STUDENT WILL"	"RESULTING IN"
CRITERIA	*Information you plan to give to students:* *-Instructions about what to do* *-Question or problem* *··Criterion input(s) to be presented (new or old)* *··Aids to be presented (if any)* *AND* *-If recognition test is to be used:* *··Answer options (i.e., actions)*	*What students are expected to do:* *-Action(s) to be taken (new or old)* *··Mode of action (recognition, editing, production)*	*What students are expected to turn out:* *-Outputs to be produced* *··Answers* *··Products (tangible, non-tangible)* *-Type of scoring* *-Standards for scoring above*

F.2.3

IDENTIFICATION MATRIX

CRITERIA FOR IDENTIFYING TEST ITEMS OR TEST PROCEDURES INVOLVING THREE RESPONSE MODES

TYPES OF TEST ITEMS	RECOGNITION	EDITING	PRODUCTION
CRITERIA	*The "GIVEN" contains:* *-A statement of a question or problem which includes:* *··Criterion INPUT(S)* *AND* *-Answer options (the criterion action associated with the input plus false choices)* *The "STUDENT WILL" requires the student to:* *-Select from among the options·* *··Checking correct option (i.e., ACTION associated with the INPUT)*	*The "GIVEN" contains:* *-A statement of a question or problem which includes:* *··Criterion INPUT(S)* *AND* *-ACTION associated with it:* *··Wrong action (to be corrected)* *··Correct action (to be left unchanged)* *The "STUDENT WILL" requires the student to:* *-Edit or correct an incorrect action* *-Leave a correct action alone*	*The "GIVEN" contains:* *-A statement of a question or problem which includes:* *··Criterion INPUT(S)* *The "STUDENT WILL" requires the student to:* *-Produce the ACTION associated with the INPUT*
EXAMPLES	*e.g., which of the following verbs do you use when the subject of a sentence is "it"?* *__am __is __are*	*e.g., correct this sentence if the verb is incorrect* *"It are blue."*	*e.g., make up a sentence using the words "it" and the present tense of the verb "to be"*

F.2.3

IDENTIFICATION MATRIX

CRITERIA FOR IDENTIFYING DESIRABLE PROPERTIES OF TEST ITEMS

PROPERTIES	CONTENT of test items	FORM of test items
CRITERIA	-*Should reflect the difficulties involved in learning and performing the criterion behavior; i.e., test items should sample difficult or error-prone:* ••*Discriminations* ••*Generalizations* ••*Associations* ••*Chains* -*Recognition and editing items should:* ••*Concentrate on error-prone behaviors* ••*Concentrate on behaviors involving learning difficulties*	-*Written test items should:* ••*Be concise, not wordy* ••*Be clear, unambiguous* ••*Not be tricky* ••*Provide adequate instructions* -*Recognition items should:* ••*Have problem completely stated in the stem (not in the options)* ••*Provide plausible options* ••*Preferably provide four options (to minimize chance factors in getting answer)*

TYPES OF SCORING PROCEDURES

	page
Conditions making for objectivity in scoring	60
What to do to make scoring more objective	61

F.2.3

CRITERIA FOR IDENTIFYING CONDITIONS THAT MAKE SCORING OF TEST ITEMS SUBJECTIVE OR OBJECTIVE

IDENTIFICATION MATRIX

SCORING	SUBJECTIVE when	OBJECTIVE when
CRITERIA	*Evaluation of outputs is more <u>difficult</u> because:* *-Comparison with standard output poses discrimination problems:* *··No model output is available* *··Properties of standard are not identified* *··The student-produced output is transient* *-Evaluation task creates a <u>more</u> complicated discrimination problem because it requires:* *··Judgment about degree* *··Estimating or rating* *··Measurement of properties unaided by measuring instruments* *··Measurement of psychological properties*	*Evaluation of outputs is <u>easier</u> because:* *-Comparison with standard output poses <u>lesser</u> discrimination problems:* *··A model output <u>is</u> available* *··Properties of standard <u>are</u> identified* *··The student-produced output is permanent* *-Evaluation task creates a <u>less</u> complicated discrimination problem because it requires:* *··Judgment about simple presence or absence* *··Counting* *··Measurement of properties aided by measuring instruments* *··Measurement of physical properties*
EXAMPLES	e.g., an essay e.g., an interpretation of an issue in history e.g., the quality of movements in sports	e.g., answer to a math problem e.g., spelling of a word e.g., punctuation of a sentence e.g., chemical distillate which weighs a precise amount

F.2.3

DETERMINING HOW TO MAKE SCORING MORE OBJECTIVE

DECISION MATRIX

CONDITIONS	Discriminations involved in evaluation are likely to be more difficult because of subjectivity concerning standards	Discriminations involved in evaluation are likely to be more difficult because of subjectivity in the evaluation task
ACTION TO TAKE	*-Produce a standard for comparison whenever possible, and/or* *-Identify the properties of a standard output:* *••A list of required properties* *-Create a permanent, tangible record of the student output, if possible; e.g.,* *••A film of motor behavior* *••An audio tape of speech*	*-Strive to evaluate on the basis of simple presence or absence (when possible)* *-Provide bench marks for degrees when it is necessary to:* *••Rate* *••Estimate* *-Identify an objective, behavioral output that represents the psychological property being evaluated*

JOB PROCEDURES

	page
Form to use in preparing test items	65, 66
Information to review for preparation of test items	67
Summary of procedures	68
Standards for test items	69

Form F.2(1)

LESSON [ ] OBJECTIVE [ ] FORM FOR TEST DEVELOPMENT

"GIVEN"	INFORMATION YOU PLAN TO GIVE TO STUDENT
— Instructions — Question or Problem •• INPUTS (new/old) •• AIDS (when applicable)	
•• ANSWER OPTIONS (when applicable)	

"STUDENT WILL"	WHAT THE STUDENT IS EXPECTED TO DO
— ACTIONS (new/old) •• mode	

"RESULTING IN"	WHAT THE STUDENT IS EXPECTED TO TURN OUT
— OUTPUT •• answers •• product — Type of Scoring — Standards for Scoring	

Form F 2(1)

LESSON [ ] *OBJECTIVE* [ ] *FORM FOR TEST DEVELOPMENT*

"GIVEN"	INFORMATION YOU PLAN TO GIVE TO STUDENT
— Instructions — Question or Problem •• INPUTS (new/old) •• AIDS (when applicable)	
•• ANSWER OPTIONS (when applicable)	

"STUDENT WILL"	WHAT THE STUDENT IS EXPECTED TO DO
— ACTIONS (new/old) •• mode	

"RESULTING IN"	WHAT THE STUDENT IS EXPECTED TO TURN OUT
— OUTPUT •• answers •• product — Type of Scoring — Standards for Scoring	

F.2.3

DECISION MATRIX

DETERMINING WHAT INFORMATION TO REVIEW IN PREPARING CRITERION TEST ITEMS			
INFORMATION NEEDED	What to include in a criterion test item: ••GIVEN ••STUDENT WILL ••RESULTING IN	How much of criterion behavior to sample and with how many items	Whether to test for criterion behavior or for simulation of criterion behavior
ACTION TO TAKE	*Review:* *-Statement of criterion objective on FORM D.2(1)* *and, if necessary,* *-Associated task analysis results on collection of FORMS A.5(4) or (11)*	*Review:* *-Statement of criterion objective on FORM D.2(1)* *and* *-Decision made in Sub-STEP F.2.2*	*Review:* *-Simulation decision made on Sub-STEP E.2.3 (based on alternate plans recorded on FORM E.1(1)*

F.2.3

ILLUSTRATION OF PROCEDURES INVOLVED IN PREPARING TEST ITEMS

#1

REVIEW

a. Statement of criterion objectives on FORM D.2(1) for three elements of a test item

b. Decision about sampling of criterion behavior (Sub-STEP F.2.2)

c. Decision about whether to use simulation (in instruction and in testing) of criterion behavior (and how to simulate) on FORM E.1(1)

#2

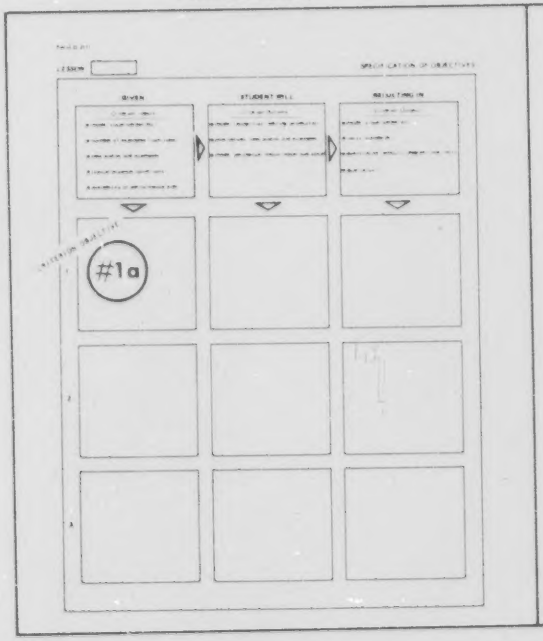
DEVELOP CRITERION TEST ITEMS

a. Spell out the "GIVEN"

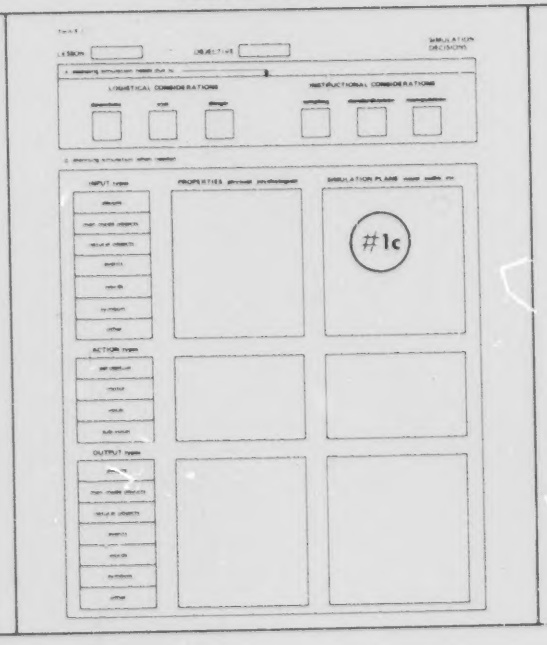
b. Describe the "STUDENT WILL"

c. Provide the "RESULTING IN" including standards

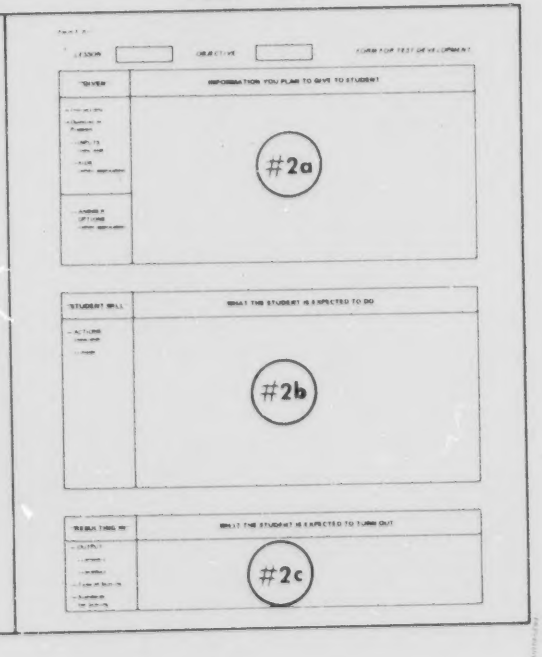
FORM D.2(1)



FORM E.1(1)



FORM F.2(1)



F.2.3

CRITERIA FOR ASSESSING THE ADEQUACY OF CRITERION TEST ITEMS

STANDARDS MATRIX

PROPERTIES	CONTENT COMPLETENESS	REFLECTS PRIOR DECISIONS	FORMAL PROPERTIES
CRITERIA	-*Three elements prepared* ••*GIVEN* ••*STUDENT WILL* ••*RESULTING IN* -*Content for three elements complete* [*See FORM F.2(1)*]	-*Statement of criterion objectives* -*Simulation decision* -*Sampling decision*	-*Lends itself to objective scoring* -*Brief, concise* -*Clear, unambiguous*

PREVIEW OF THE NEXT SubSTEP

YOUR PRODUCT	*A test containing items testing for subcriterion objectives.*
WHAT YOU WILL WORK FROM	(1) Statements of subcriterion objectives. (2) Decision about how much behavior to sample and how many items to develop. (3) Decision whether or not to simulate.
WHAT YOU WILL DO	(1) Develop subcriterion test items.
FORMS YOU WILL USE	FORM F.2(1) for developing test items.

DESCRIPTION OF Sub-STEP	F.2.4

INPUT	ACTION	OUTPUT
Statement of subcriterion objective + sampling decision + simulation decision x	Prepare subcriterion test item(s) xi	SUBCRITERION test item(s) containing: -Giv -STUDENT WILL -RESULTING IN xii

Job Aid Contents

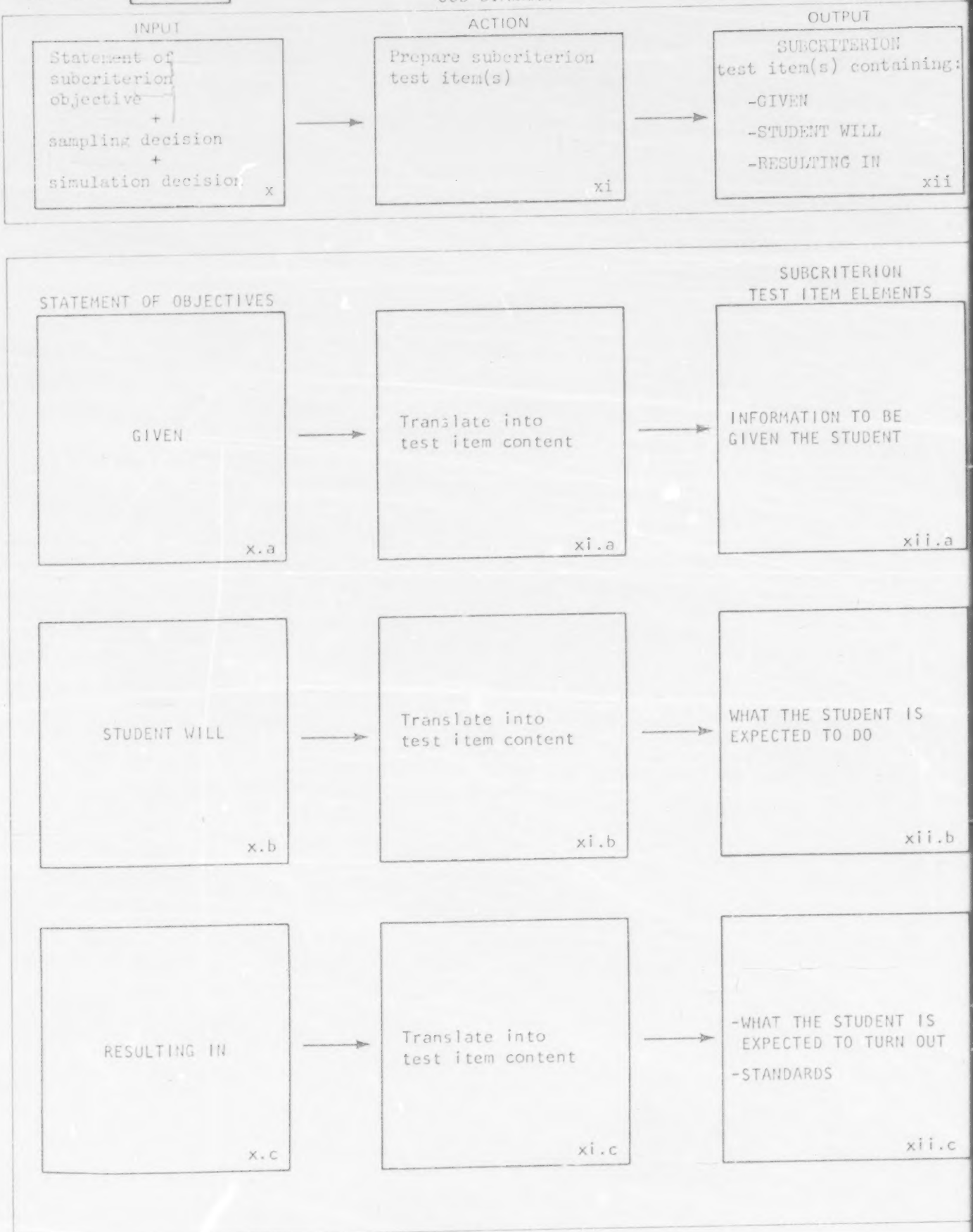
CRITERIA FOR IDENTIFYING INPUTS	ACTION TO BE TAKEN	STANDARD FOR OUTPUTS	FORMS TO USE
-MATRIX: Difference between criterion and subcriterion test item 74			FORM F.2(1) SUMMARY OF PROCEDURES . . 75

Required Materials

COMPLETED MATERIALS	STEP	COMPLETED FORMS	STEP	BLANK FORMS
Sampling decision	F.2.2	Collection of FORMS carried forward from	F.2.2	FORM F.2(1)
Selection of simulation plan (where applicable)	E.2.3			

Sub-STEP F.2.4

JOB DIAGRAM



JOB PROCEDURES

	page
Difference between criterion and subcriterion test items	74
SUMMARY OF PROCEDURES	75

F.2.4

IDENTIFICATION MATRIX

CRITERIA FOR DISTINGUISHING BETWEEN CRITERION TEST ITEMS AND (UNMODIFIED) PREPARATORY TEST ITEMS

TYPE OF TEST ITEM	CRITERION test item	SUBCRITERION test item
CRITERIA	*The test item tests for the entire chain involved in criterion behavior**	*The test item tests for only part of the chain; but it is an intact part, i.e., the criterion behavior is not modified (except for scope) in any way*
NOTE: Each item tests for only one of the many possible variations a criterion chain can have; but it still tests an entire chain.		
IMPLICATIONS	Considerations in developing test items for subcriterion behaviors, except for scope, are identical with those involved in preparing criterion test items (See Sub-STEP F.2.3)	

F.2.4

ILLUSTRATION SUMMARIZING PROCEDURES INVOLVED IN PREPARING SUBCRITERION TEST ITEMS (REGARDING UNMODIFIED PARTS OF CRITERION BEHAVIOR)

#1

REVIEW

a. Statement of subcriterion objectives on FORM D.2(1) for three elements of a test item

b. Decision about sampling of subcriterion behavior (Sub-STEP F.2.2)

c. Decision about whether to use simulation (in instruction and in testing) of subcriterion behavior (and how to simulate) on Form E.1(1)

#2

DEVELOP SUBCRITERION TEST ITEMS

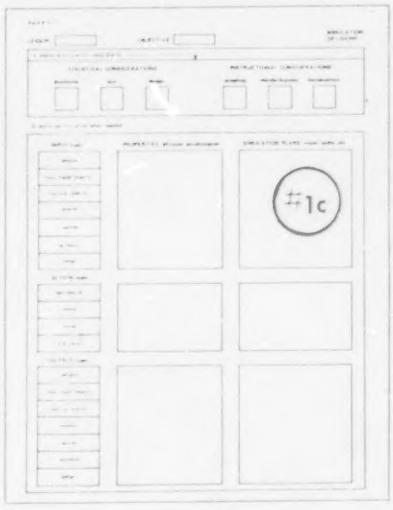
a. Spell out the "GIVEN"

b. Describe the "STUDENT WILL"

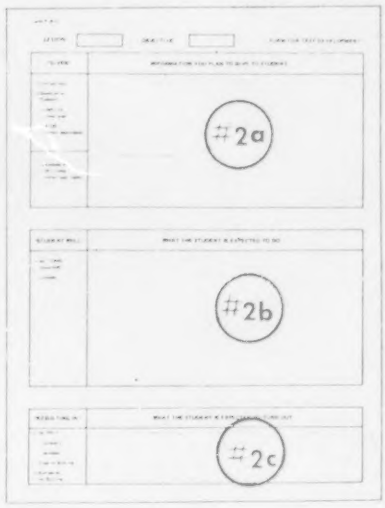
c. Provide the "RESULTING IN" including standards

FORM D.2(1)

FORM E.1(1)



FORM F.2(1)



STEP F.2

COMPLETION CHECKLIST

	IDENTIFIED	PERFORMED	PRODUCED	FORMS COMPLETED
F.2.1		Decision made whether to assess criterion and/or subcriterion outputs		
F.2.2		Decision made about how much of criterion behavior to sample and how many items to prepare		
F.2.3			Criterion test items	FORM F.2(1)
F.2.4			Subcriterion test items	FORM F.2(1)

STEP F.3

F.3 Develop diagnostic tests to identify reason for failure to acquire the criterion behavior taught in each instructional unit.

F.3.1 Plan the use of results on criterion ~~test~~ items and on subcriterion test items in diagnosing type of learning failure that has occurred.

F.3.2 Develop diagnostic interview procedures to provide supplementary diagnostic information about results on criterion test items and on subcriterion test items

F.3.3 Develop diagnostic test items designed to identify specific types of learning failures.

STEP F.3

OVERVIEW

INPUT	ACTION	OUTPUT
Statement of objectives and associated analysis + criterion and subcriterion test items for each criterion behavior	Plan the way to diagnose learning failures regarding the criterion behavior	Plan involves: -Use of criterion tests -Use of subcriterion tests -Development of diagnostic items

F.3.1	Subcriterion objectives associated with a criterion behavior + developed criterion tests and preparatory tests i	Review to determine adequacy of criterion and subcriterion tests to serve a diagnostic purpose ii	DECISION -Criterion tests and subcriterion tests can diagnose problems AND/OR -Diagnostic tests are needed iii
F.3.2	Statement of objectives + task analysis results + learning analysis results + criterion test items iv	Develop diagnostic probes when necessary v	PROBES To diagnose type of learning failure and sources of the failure vi
F.3.3	Statement of objectives + task analysis results + learning analysis results + criterion test items vii	Develop diagnostic test items appropriate to anticipated types of learning failures viii	DIAGNOSTIC TEST ITEM to diagnose type of learning failure ix

STEP F.3

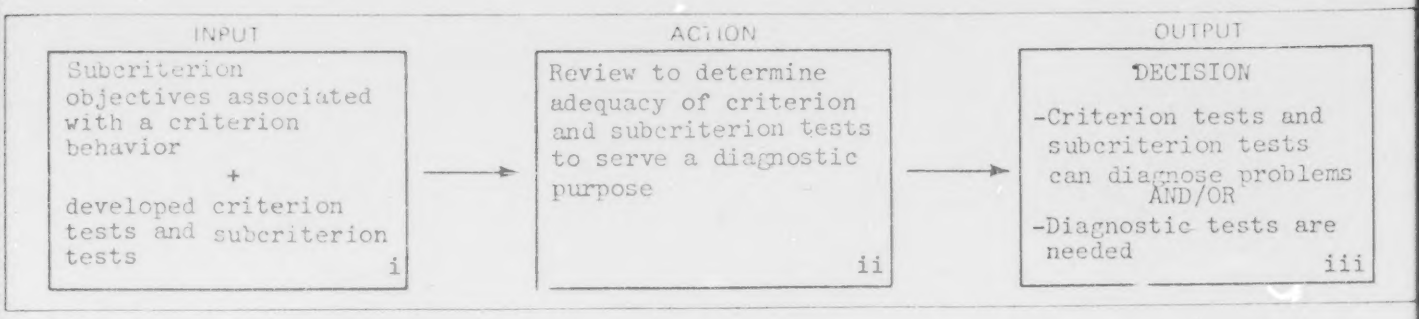
PAGE INDEX

	CRITERIA FOR IDENTIFYING INPUTS	ACTION TO BE TAKEN	STANDARD FOR OUTPUTS	FORMS TO USE
F.3.1	-MATRIX: Non-diagnostic nature of single item results . . . 86-88	-MATRIX: Using multiple test items to diagnose learning failures . . . 91-103 -MATRIX: Information to review in planning diagnosis . . . 106		SUMMARY OF PROCEDURES . . .107
F.3.2	-MATRIX: Using probes regarding criterion tests vs. diagnostic tests 111	-MATRIX: Types of probes to use . . . 116-121 -MATRIX: Developing probes . . 123-128	-MATRIX: Adequacy of probes which have been developed . . . 129	SUMMARY OF PROCEDURES . . 128
F.3.3	-MATRIX: Types of diagnostic test items . . . 133-137	-MATRIX: Priorities in order of administering test items . . . 140 -MATRIX: Information to review as a basis for developing diagnostic test items . . . 141	-MATRIX: Adequacy of diagnostic test items . . 143	F.2(1) SUMMARY OF PROCEDURES . . 142

PREVIEW OF THE NEXT SubSTEP

YOUR PRODUCT	*Determining whether tests are capable of diagnosing the types of learning failures which have occurred.*
WHAT YOU WILL WORK FROM	(1) Subcriterion and criterion objectives. (2) Subcriterion and criterion tests.
WHAT YOU WILL DO	(1) Review test items to determine whether anticipated learning failures can be diagnosed by them. (2) Decide whether the existing tests are adequate or inadequate for diagnostic purposes.
FORMS YOU WILL USE	None

DESCRIPTION OF Sub-STEP	F.3.1



Job Aid Contents

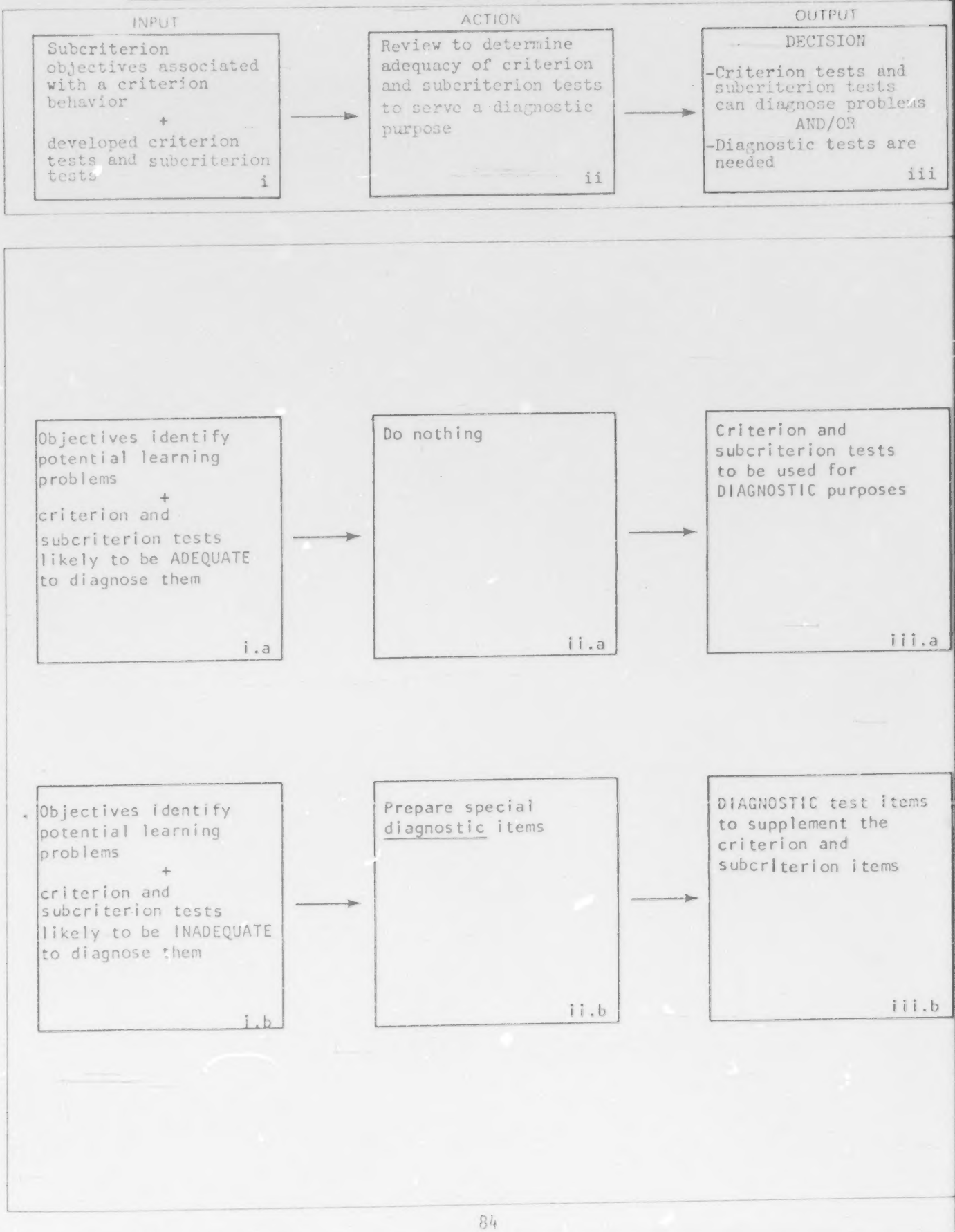
CRITERIA FOR IDENTIFYING INPUTS	ACTION TO BE TAKEN	STANDARD FOR OUTPUTS	FORMS TO USE
-MATRIX: Non-diagnostic nature of single item results . . . 86-88	-MATRIX: Using multiple test items to diagnose learning failures . . 91-103 -MATRIX: Information to review in planning diagnosis . . . 106		SUMMARY OF PROCEDURES . . . 107

Required Materials

COMPLETED MATERIALS	STEP	COMPLETED FORMS	STEP	BLANK FORMS
Analysis materials and statement of objectives (carried forward)	F.2	Criterion and subcriterion tests	F.2.4	

Sub- STEP F.3.1

JOB DIAGRAM



INABILITY OF USING RESULTS
ON SINGLE CRITERION OR SUBCRITERION TEST ITEMS
TO DIAGNOSE LEARNING FAILURES

	page
Type of diagnostic information you wish to obtain	86
How the same wrong answer may be due to a variety of learning failures	88

F.3.1

CRITERIA FOR IDENTIFYING TYPES OF DIAGNOSTIC INFORMATION YOU WISH TO OBTAIN: WHAT IS THE SOURCE OF FAILURE

IDENTIFICATION MATRIX

SOURCES OF FAILURE	Re: INPUTS -Discriminations -Generalizations	Re: ACTIONS -Generalizations -Associations -Chains
CRITERIA	*Has the student* *-Treated different classes of inputs in the same way when he should have:* *••Discriminated among the classes* *AND* *••Taken different actions in response to them* *-Treated members belonging to the same class of inputs in different ways when he should have:* *••Generalized across the members* *AND* *••Taken the same action in response to them*	*Has the student* *-Failed at action generalization:* *••Failed to exhibit an alternative way to take an action* *••Taken an incorrect alternative action* *-Confused which action goes with which class of inputs when he should have:* *••Properly associated the right input with right action* *-Failed to complete a chain or failed to perform its parts in the correct sequence*

EXAMPLES	See opposite page

F.3.1

EXAMPLES OF LEARNING FAILURES

EXAMPLES	Re: INPUTS	Re: ACTIONS
ENGLISH *noun/verb agreement*	A DISCRIMINATION failure e.g., *student writes "The women is sick"* -Failure to discriminate between *"woman" and "women"* belonging to the two <u>different</u> classes of inputs: *singular and plural* -Has treated the two classes as if they were the same A GENERALIZATION failure e.g., *student writes "The women is sick"* -Fails to generalize across *regular and irregular plurals, all of which belong to the <u>same</u> class* -Has treated members of the same class as if they were different	An ASSOCIATION failure e.g., *student writes "The women is sick"* -Failure to associate *plural nouns and the plural form of the verb* -Has associated the wrong action with the input presented to him
NOTE:	The same error, *i.e., writing "is,"* can be indicative of three different types of learning failure. By itself, the result is <u>not</u> diagnostic. It is indeterminate which type of error has been made. See discussion on next page.	

F.3.1

IDENTIFICATION MATRIX

CRITERIA FOR IDENTIFYING THE NON-DIAGNOSTIC NATURE OF RESULTS ON SINGLE CRITERION TEST ITEMS

TYPES OF FAILURE	DISCRIMINATION among inputs	GENERALIZATION across inputs	ASSOCIATION between input and action
CRITERIA	*On a criterion test item all the student is required to do is to produce a criterion ACTION. Any of the three types of failure will show up as a wrong (or omitted) action. The same wrong action can be due to any of the three failures. Therefore, its source is indeterminate.*		

EXAMPLE*	A student is given the following test item: *A stress has been removed from this piece of paper (it is now crumbled). Is this a perfectly elastic or a non-perfectly elastic object?* The student gives the wrong answer; takes the wrong ACTION: *Says it is "perfectly elastic."* The failure may be due to an inability: -To discriminate between examples of *perfectly elastic and non-perfectly elastic objects* OR
	-To generalize across varying (dissimilar) examples of *non-perfectly elastic objects* OR -To associate the label *"non-perfectly elastic" with an example*

*See previous page for another example.

HOW TO USE MULTIPLE
CRITERION OR PREPARATORY TEST ITEMS
FOR DIAGNOSTIC PURPOSES

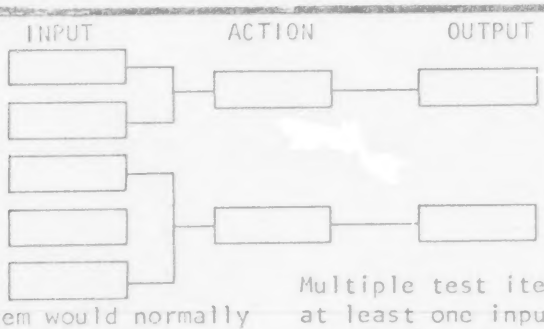
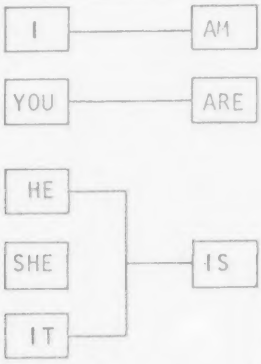
Problems re: INPUTS ••Discriminations ••Generalizations / INPUTS AND ACTIONS ••Associations	page
How multiple test items sample more varieties of the criterion behavior	91
Properties of test items and results which can be used diagnostically	92
Interpreting consistency vs. inconsistency in error patterns	94, 95
Interpreting some general patterns of results*	96

*The number of combinations and permutations (of: right and wrong answers, types of wrong answers, and kinds of items) is too large to be able to identify them all and to interpret them unequivocally. This section and related sections in Part J on "the evaluation of instructional materials," therefore, only identify patterns of results which lend themselves to interpretations which can be made with reasonable certainty.

F.3.1

CRITERIA FOR IDENTIFYING WHAT TYPE OF "MULTIPLES" ADDITIONAL TEST ITEMS SAMPLE

IDENTIFICATION MATRIX

NUMBER OF TEST ITEMS	A SINGLE test item	MULTIPLE test items
CRITERIA	*-When there are multiple classes of INPUTS or of ACTIONS, and/or* *-When there are multiple members within each class,* *A single test item usually samples:* *··Only one class, or* *··Only one member within a class*	*A multiple test item usually samples:* *··More than one class, or* *··More than one member of a class*
GENERAL EXAMPLE	 A single test item would normally present only one of the five diagrammed INPUTS	Multiple test items might present at least one input for each of the two diagrammed classes of inputs, and possibly two inputs per class
EXAMPLE	 A single test item might present the student with the pronoun "she" (and require a sentence with "is" in it)	Multiple test items are likely to present all five pronouns as INPUTS (with the student required to produce the associated ACTION, i.e., using the right verb)

F.3.1

CRITERIA FOR IDENTIFYING PROPERTIES OF CRITERION (OR SUBCRITERION) TESTS WHICH HAVE DIAGNOSTIC CAPABILITIES

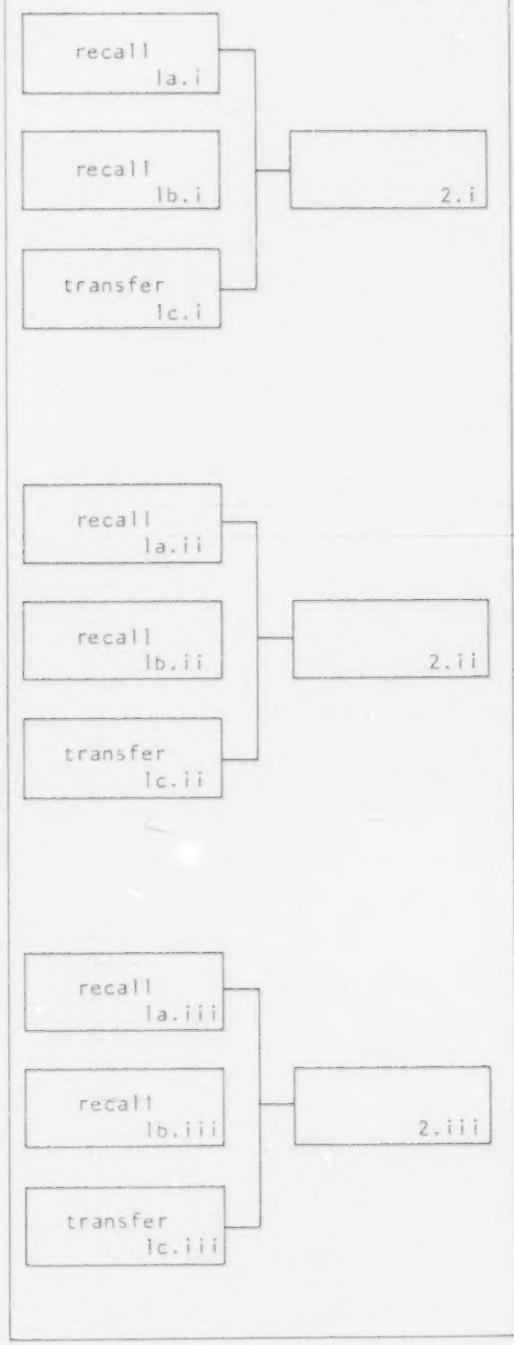
IDENTIFICATION MATRIX

TYPES OF PROPERTIES	Independent variables manipulated by you	Dependent variables measured by you
CRITERIA	-*Types of items presented to students* ••*Recall* ••*Transfer* -*Number of items presented to students* ••*Number of classes (inputs/actions) each represented by a test item* ••*Number of members per class each represented by a test item* ••*Number of test items per class (input/action)* ••*Number of test items per each member of a class* -*Reversal in the direction of performance required by test items* ••*Criterion INPUT is the "given" in the test item; criterion ACTION is what the "student will" do* ••*The same criterion ACTION becomes the "given" in a reversal item, and the criterion INPUT is what the "student will" produce*	-*Nature of errors made by students* ••*Taking the same action for inputs requiring different actions* ••*Taking different actions for inputs requiring the same action* ••*Mismatching input and action* ••*Actions are omitted altogether* -*Consistency of errors made* ••*The same/different type of error made on recall and on transfer items which are parallel to one another, e.g., represent the same INPUT class* ••*The same/different type of error made on more than one item per class or per member of a class* ••*The same/different error made in a reversal item (if used)*

EXAMPLES	SEE OPPOSITE PAGE

EXAMPLES ILLUSTRATING PROPERTIES OF TEST ITEM (RESULTS)
WHICH CAN BE USED DIAGNOSTICALLY

EXAMPLES



-Nature of errors made by students

e.g., GIVEN: As inputs on separate test items: 1a.i, 1b.ii, and 1c.i

STUDENT: Takes the same actions (2.ii) for all three inputs instead of taking three different actions

---

e.g., GIVEN: As inputs on separate test items: 1a.i, 1a.ii, and 1a.iii

STUDENT: Takes different actions (1.i, 2.ii, and 2.iii) for each of them instead of taking the same action (2.i) for them all

---

e.g., GIVEN: As inputs on separate test items: 1a.i, 1a.ii, and 1a.iii

STUDENT: Takes three different actions (2.i, 2.ii, and 2.iii) but matches them up with the wrong inputs (1a.i with 1.iii), (1a.ii with 2.i), and (1a.iii with 2.ii)

---

e.g., GIVEN: As inputs on separate test items--a recall item (1a.i) and a transfer item (1a.iii)

STUDENT: A student makes an error on one (the wrong action--2.ii) but not the other

---

e.g., GIVEN: As inputs on separate items two instances of the same member of a class (1b.ii)

STUDENT: Takes the same wrong action (2.ii) for both of them

F.1.3

IDENTIFICATION MATRIX

CRITERIA FOR IDENTIFYING CONSISTENCY AND INCONSISTENCY IN ERROR PATTERNS ON MULTIPLE TEST ITEMS

ERROR PATTERN	CONSISTENT errors	INCONSISTENT errors
CRITERIA	*On multiple test items:* *-The same wrong action is taken to successive (multiple) examples of INPUTS from the same class*	*On multiple test items:* *-A different action is taken to successive (multiple) examples of INPUTS from the same class:* *··Right and wrong actions may be found* *··Different wrong actions may be found*
EXAMPLES INPUT → ACTION I → am you → are he, she, it → is	*e.g., when given the pronoun "you" on successive test items,* *the student consistently uses the verb form "is"*	*e.g., when given the pronoun "you" on successive test items,* *the student is inconsistent; sometimes he uses "are," and sometimes he uses "is"*

F.3.1

DECISION MATRIX

DETERMINING HOW TO INTERPRET CONSISTENCY/INCONSISTENCY ON MULTIPLE TEST ITEMS

ERROR PATTERN	CONSISTENT errors	INCONSISTENT errors
ACTION TO TAKE	*INTERPRET consistency in error pattern as:* *-Additional confirmation of diagnosis* *-Evidence of a <u>sustematic</u> error; i.e., a <u>definite</u> incorrect pattern has been learned*	*INTERPRET inconsistency in error pattern as:* *-Evidence of a variable error; i.e., <u>no</u> definite pattern has been learned*

EXAMPLES	REVIEW EXAMPLE ON PREVIOUS PAGE	
	-In view of the fact that the student consistently says "is" when given "you" as a subject, *it is concluded that he has systematically associated that form of the verb with the pronoun "you"*	*-In view of the fact that the student sometimes says "is" and sometimes says "are,"* *it is concluded that his error is variable and not systematic. He <u>has not</u> associated a particular form of the verb with the pronoun "you"*

F.3.1

DETERMINING HOW TO INTERPRET GENERAL ERROR PATTERNS REGARDING DISCRIMINATIONS, GENERALIZATIONS, AND ASSOCIATIONS

DECISION MATRIX

ERROR PATTERNS	-All INPUT classes are treated differently, i.e., a different ACTION is taken in response to them; BUT INPUTS AND ACTIONS are mismatched	-The same ACTION is taken to two or more INPUT classes	-A different ACTION is taken in response to some members of the same class	-Recall test items are correct -Transfer test items are incorrect
ACTION TO TAKE *"Make these general interpretations"*	*INTERPRET as follows:* *-Since all INPUT classes are treated differently, the classes have been correctly discriminated* *-Consider this a problem in association between INPUT and ACTION*	*INTERPRET as follows:* *-Since two or more INPUT classes have been treated the same way, they have not been correctly discriminated* *-Consider this a problem in discrimination among classes of INPUTS*	*INTERPRET as follows:* *-Since members of the same class have been treated in different ways when they should have been treated alike* *-Consider this a problem in generalization across INPUTS within a class*	*INTERPRET as follows:* *-Since the recall items are correct, it means that discriminations and associations have been correctly learned* *-Consider this a generalization problem*

EXAMPLES	SEE OPPOSITE PAGE

F.3.1

EXAMPLES ILLUSTRATING INTERPRETATIONS OF SOME GENERAL ERROR PATTERNS

EXAMPLES

ERROR PATTERN	-All INPUT classes are treated differently--BUT -INPUT classes and ACTIONS are mismatched Problem: ASSOCIATION	-The same ACTION is taken for two or more classes Problem: DISCRIMINATION	-A different ACTION is taken for some members belonging to the same class Problem: GENERALIZATION	-Recall test items O.K. -Transfer test items not O.K. Problem: GENERALIZATION
EXAMPLES Singular nouns → Verbs in present tense with "s" ending Plural nouns → Verbs in present tense without "s" ending	TEST RESULTS: -All singular nouns (one class of inputs) are responded to with the plural form of the verb -All plural nouns (the second class of inputs) are responded to with the singular form of the verb INTERPRETATION: -Singular and plural nouns are successfully discriminated -All singulars are treated alike; all plurals are treated alike. Generalization is therefore successful -The two classes of inputs and the actions are mismatched: An association problem	TEST RESULTS: -Both singular and plural nouns are responded to with verbs having no "s" ending; e.g., "The boy walk to work" INTERPRETATION: -The two classes of INPUTS (singular and plural) are not successfully discriminated	TEST RESULTS: -Some plural nouns are responded to correctly and some incorrectly (i.e., with no "s" and with an "s" ending" INTERPRETATION: -Not all members of the class are seen as plurals (e.g., irregular plurals) -A generalization problem	TEST RESULTS: -All test items using nouns that appeared in instruction (RECALL items) are responded to correctly -TRANSFER items (new nouns not used in instruction) are responded to incorrectly INTERPRETATION: -Discrimination between singular and plural has been successfully learned -Association between noun form and verb form has been successfully learned -Generalization to other singulars or to other plurals has NOT been successfully learned

INTERPRETING ERROR PATTERNS
INVOLVING ACTIONS ON CRITERION TESTS

Problems regarding: ACTIONS ••Generalizations ••Chaining	page
Properties of test results which can be used diagnostically	100
Interpreting some general patterns of results	102

F.3.1

IDENTIFYING PROPERTIES OF ERROR PATTERNS INVOLVING "ACTIONS"

IDENTIFICATION MATRIX

TYPE OF PROPERTIES	ALTERNATIVE* actions	SEQUENCE* of actions
a CRITERIA	-*Nature of errors made by students* ••*Omitted taking an alternative action (one action already learned)* ••*Taking an incorrect alternative action* -*Consistency of errors made* ••*Repetition of error on retesting*	-*Nature of errors made by students* ••*Omitted exhibiting the entire series of associations* ••*Omitted exhibiting part of the series* ••*Exhibited parts of the series out of sequence* -*Consistency of errors made* ••*Repetition of error on retesting*

*These are not mutually exclusive categories; alternative actions may involve sequences, and there may be alternative sequences.

EXAMPLES	SEE OPPOSITE PAGE

F.3.1

EXAMPLES CONCERNING ERROR PATTERNS INVOLVING "ACTIONS"

EXAMPLES

TYPE OF PROPERTIES	ALTERNATIVE* actions	SEQUENCE* of actions
EXAMPLES	PROBLEMS REGARDING ACTION GENERALIZATION	PROBLEMS REGARDING CHAINING
	e.g., *a student can give a verbal definition of Bernoulli's principle; but he cannot give an example of it (or vice versa)*	e.g., *in threading a film projector, a student may omit a loop altogether; or he may make the loop out of sequence (and foul up the threading)*
	e.g., *a student can repeat verbatim a learned definition of the concept "anthropomorphism"; but he cannot paraphrase the definition or cannot identify (e.g., in a list) the essential properties of the concept*	e.g., *in factoring the expression $(5-3)^2$, the student may first square each number in the parentheses rather than first subtracting and then squaring; the sequence is incorrect*
	e.g., *a student can compute a "mean" for a distribution of scores using a formula for machine computation; but he cannot compute a "mean" by hand*	e.g., *in reciting a poem or in acting in a play, the student may produce passages out of sequence*

F.3.1

DETERMINING HOW TO INTERPRET
GENERAL ERROR PATTERNS CONCERNING "ACTIONS"

DECISION MATRIX

ERROR PATTERNS	-Error patterns reveal no discrimination or generalization problem regarding INPUTS, nor an association problem between INPUT and ACTION -On a test item involving an example (new or old) from the same class of inputs, an alternative action is omitted or an incorrect alternative is taken	-Error patterns reveal no discrimination or generalization problem regarding INPUTS, nor an association problem between INPUT and ACTION -On a test item involving an example (new or old) from the same class of inputs, ••All or part of a series of associations is omitted ••The sequence of associations is incorrect
ACTION TO TAKE	*INTERPRET as follows* *-Since it has been established that INPUTS and ACTIONS have been correctly associated,* *-Consider this a problem of generalization across ACTIONS*	*INTERPRET as follows* *-Since it has been established that INPUTS and ACTIONS have been correctly associated in all the separate associations,* *-Consider this a problem in chaining a series of associations*

NOTE

More detail on interpretation of error patterns, both on practice items during instruction and on criterion/subcriterion tests, are provided in Section J: EVALUATE INSTRUCTIONAL MATERIALS.

JOB PROCEDURES

	page
Information to review in planning the use of criterion or subcriterion test items for diagnostic purposes	106
SUMMARY OF PROCEDURES	107

F.3.1

DECISION MATRIX

DETERMINING WHAT INFORMATION TO REVIEW IN PLANNING THE USE OF CRITERION OR PREPARATORY TESTS FOR DIAGNOSTIC PURPOSES

INFORMATION NEEDED	For each CRITERION BEHAVIOR are there likely to be learning failures involved in learning some part or all of the criterion behavior?	Are there likely to be -Enough criterion and subcriterion test items -Varied enough criterion and subcriterion test items to diagnose precise type of learning failure?
ACTION TO TAKE	*-Review the set of "statement of objectives" associated with each criterion behavior for:* *••Preparatory objectives identifying component skills which are likely to pose learning difficulties*	*-Review the entire set of criterion and subcriterion test items for coverage of:* *••All the INPUT classes and ACTION classes likely to be involved in learning failures* *••Recall/transfer items relating to the INPUT/ACTION classes* *••Multiple examples within each class* *••Multiple instances of the same example (a retest item)*

F.3.1

ILLUSTRATION SUMMARIZING PROCEDURES INVOLVED IN PLANNING THE USE OF CRITERION OR PREPARATORY TEST ITEMS FOR DIAGNOSTIC PURPOSES

#1

For EACH CRITERION BEHAVIOR REVIEW:

a. The subcriterion objective which is associated with it and which identifies potential learning problems

b. The criterion and subcriterion tests for adequacy of coverage of potential learning problems (and likely failures)

#2

DECIDE

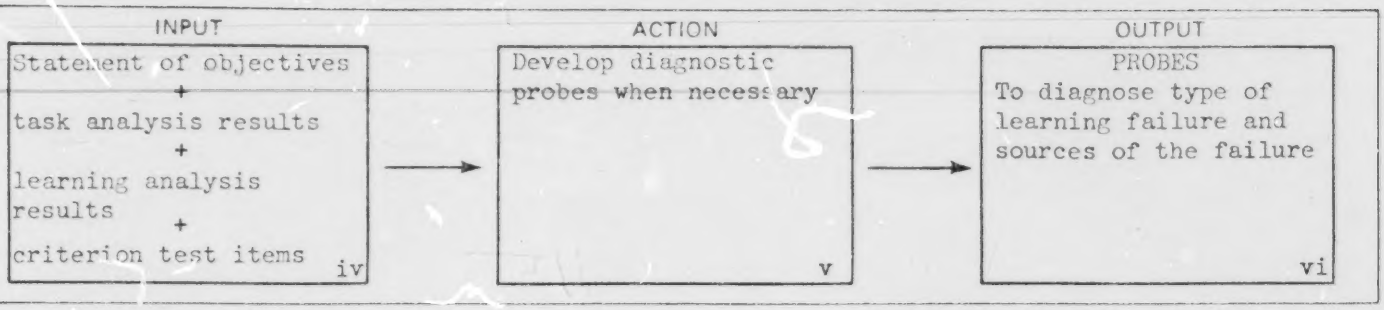
a.1 To prepare diagnostic items--when coverage by criterion or subcriterion items is insufficient for diagnostic purposes

a.2 Not to prepare diagnostic items--when coverage by criterion or subcriterion items is sufficient for diagnostic purposes

PREVIEW OF THE NEXT SubSTEP

YOUR PRODUCT	*Diagnostic probes to supplement test procedures as a means of diagnosing learning failures and their probable causes.*
WHAT YOU WILL WORK FROM	(1) Statements of objectives (2) Task and learning analyses indicating what the likelihood is of learning failures occurring (3) Criterion test items
WHAT YOU WILL DO	(1) Review above information and determine whether tests are likely to be sufficiently diagnostic of learning failures and the causes for them. (2) Develop diagnostic probes when necessary.
FORMS YOU WILL USE	None

DESCRIPTION OF Sub-STEP	F.3.2



Job Aid Contents

CRITERIA FOR IDENTIFYING INPUTS	ACTION TO BE TAKEN	STANDARD FOR OUTPUTS	FORMS TO USE
-MATRIX: Using probes regarding criterion tests vs. diagnostic tests 111	-MATRIX: Types of probes to use . . . 116-121 -MATRIX: Developing probes . . 123-128	-MATRIX: Adequacy of probes which have been developed . . . 129	SUMMARY OF PROCEDURES . . 128

Required Materials

COMPLETED MATERIALS	STEP	COMPLETED FORMS	STEP	BLANK FORMS
Determination of likely need of diagnostic information	F.3.1	Forms associated with criterion behavior (carried forward from)	F.3.1	

Sub-STEP F.3.2

JOB DIAGRAM

INPUT	ACTION	OUTPUT
Statement of objectives + task analysis results + learning analysis results + criterion test items (iv)	Develop diagnostic probes when necessary (v)	PROBES To diagnose type of learning failure and sources of the failure (vi)

Anticipated DISCRIMINATION problems (iv.a)	Develop specific probes to determine this type of problem and its source (v.a)	QUESTIONS -Re: INPUTS and their properties -Re: INPUT examples requiring indication of "same" or "different" -Re: Examples varied systematically for similarity (vi.a)
Anticipated GENERALIZATION problems (iv.b)	Develop specific probes to determine this type of problem and its source (v.b)	QUESTIONS -Re: INPUTS and their properties -Re: INPUT examples requiring indication of "same" or "different" -Re: Examples varied systematically for similarity (vi.b)
Anticipated ASSOCIATION problems (iv.c)	Develop specific probes to determine this type of problem and its source (v.c)	QUESTIONS -Re: INPUTS and ACTIONS -Reverse direction of performance -Correctness of pairings (vi.c)

TWO APPROACHES TO OBTAINING
SPECIFIC DIAGNOSTIC INFORMATION

	page
Two ways to obtain specific diagnostic information	112
When to use each of the two diagnostic approaches	113

F.3.2

IDENTIFICATION MATRIX

CRITERIA FOR IDENTIFYING TWO WAYS TO OBTAIN SPECIFIC DIAGNOSTIC INFORMATION

TWO METHODS	Preparing and using DIAGNOSTIC TEST ITEMS	Preparing and using PROBING QUESTIONS with students concerning their performance on criterion and preparatory test items
CRITERIA	-*Test items which are designed specifically to identify failures:*	-*Questions used in interviews with students following their performance (and commission of errors on) criterion and preparatory tests* -*Questions which are designed specifically to identify failures:*
	Regarding: INPUTS *••Discriminations* *••Generalizations* *Regarding: INPUTS and ACTIONS* *••Associations* *••Chains* *Regarding: ACTIONS* *••Generalizations*	

F.3.2

DETERMINING WHEN TO USE EACH OF THE TWO WAYS TO OBTAIN SPECIFIC DIAGNOSTIC INFORMATION*

DECISION MATRIX

CONDITIONS	-A full range of multiple criterion (or preparatory) test items have not been developed, thus not allowing patterns of results to be analyzed -Multiple test item patterns are difficult to interpret	-Specific preparatory objectives have been stated which identify specific component skills whose required acquisition is likely to pose learning difficulties
ACTION TO TAKE	*Plan to interview students and probe for sources of learning failure*	*Plan to develop special diagnostic tests*
METHODS	See page 115	See page 131

*Both methods are likely to be used during development of instructional materials. In the routine administration of a fully developed and validated instructional program, only diagnostic tests are likely to be used. At that time their use is concerned with the specific problems of individual students rather than with problems about the adequacy of the instructional materials.

FOLLOWING ADMINISTRATION OF CRITERION TESTS:
PROBING FOR SOURCES OF LEARNING FAILURES*

	page
Priorities in order of question types used	116
Two types of errors: (1) omissions and (2) wrong answers, requiring different interpretations and probing	117
Interpreting replies to probes indicating: (1) something not learned, OR (2) something incorrectly learned	118
Probes to determine whether: (1) something not learned, OR (2) something incorrectly learned	120
Other probing procedures to use to identify specific types of learning failures	121

Similar probing can be performed with practice items during informal tryouts of instructional materials; a one-to-one tryout and interview session involving an administrator and a single student can be used to identify sources of failures on criterion practice items which are a part of the instructional program. See SECTION J. Section J also provides detailed considerations in interpreting errors on single test items, responses to test probes, and patterns of errors on multiple test items.

F.3.2

DECISION MATRIX

DETERMINING PRIORITIES IN THE ORDER IN WHICH - DIFFERENT TYPES OF PROBING QUESTIONS SHOULD BE ASKED

PRIORITIES		If first answers provide indeterminate results	If second answers still provide indeterminate results
	FIRST	SECOND	THIRD
ACTION TO TAKE	-*Ask open-ended questions which do not identify or suggest answers he might give* -*Do not call attention either to the INPUTS or to the ACTIONS for him to focus on* -*Do not call attention to specific properties of INPUTS or of ACTIONS to focus on*	-*Ask directed questions which identify what the student should consider when trying to assess what went wrong* -*Call attention generally either to the INPUTS or to the ACTIONS for him to focus on, BUT* -*Do not call attention to specific properties of INPUTS or of ACTIONS to focus on*	-*Ask directed questions* -*Call attention to INPUTS or to ACTIONS* -*Call attention to specific properties for student to focus on*
EXAMPLES	-"Why did you _____ (take this ACTION)?" e.g., *"Why did you classify the leaf this way?"*	-"What is there about this _____ (INPUT) that made you _____ (take this ACTION)?" e.g., *"What property or attribute of this leaf made you classify it this way?"*	-"Which of these _____ (properties) of the _____ (INPUTS) made you _____ (take this ACTION)?" e.g., *"Which of these leaf properties: size, shape, or color made you make this particular classification?"*

F.3.2

DETERMINING HOW TO INTERPRET TWO TYPES OF ERRORS

DECISION MATRIX

TYPES OF ERRORS	NO action is taken on a test item ••An omission ••No answer given	A WRONG action is taken on a test item ••A wrong answer is given
ACTION TO TAKE (INTERPRET)	*On RECALL items* *-Most likely indicative of:* *••A failure to acquire or to retain a discrimination, a generalization, or an association in the first place; learning did not occur* *-Less likely (but possibly) indicative of:* *••An acquisition of an INCORRECT discrimination, generalization, or association; i.e., something wrong is learned* *on TRANSFER items* *-Most likely indicative of:* *••A failure to acquire a generalization* *-Less likely (but possibly)* *••Interpretations provided above for RECALL items*	*On RECALL items* *-Most likely indicative of:* *••An acquisition of an INCORRECT discrimination, generalization, or association; i.e., something wrong is learned* *-Less likely (but possibly) indicative of:* *••A failure to acquire or to retain a discrimination, a generalization, or an association in the first place; learning did not occur* *on TRANSFER items* *-Most likely indicative of:* *••A wrong generalization has been made; or* *••A wrong discrimination has been made* *-Less likely (but possibly)* *••Interpretations provided above for RECALL items*

F.3.2
IDENTIFICATION
MATRIX

CRITERIA FOR IDENTIFYING WHETHER LEARNING DID NOT OCCUR
OR WHETHER INCORRECT LEARNING TOOK PLACE

TYPE OF PROBLEM	Learning has NOT taken place	INCORRECT learning has taken place
CRITERIA	*HIGHLY likely when:* *-On another RECALL test item, an answer is omitted* *AND* *-In response to probes, he says one or more of the following:* *··I don't know the answer* *··I don't know why I left it out* *-In reply to probes, student does any of the following:* *··Does not attribute properties to INPUTS* *··Does not indicate that he is confused about INPUTS (or their properties)* *··Does not identify (several) possible answers*	*HIGHLY likely when:* *-On another RECALL or TRANSFER item, makes the same error* *-In reply to probes, the student does one or more of the following:* *··Attributes properties to INPUTS incorrectly* *··Indicates general or specific confusion about INPUTS (or their properties); he pays attention to irrelevant properties* *··Identifies possible answers but indicates he does not know which is the correct one*

EXAMPLES	SEE OPPOSITE PAGE

F.3.2

EXAMPLES ILLUSTRATING INTERPRETATION OF STUDENT RESPONSES TO ADDITIONAL TEST ITEMS OR TO PROBES

EXAMPLES

ORIGINAL TEST ITEM	GIVEN: An old (RECALL) example of a "first class" lever (i.e., fulcrum between effort force and resistance force) STUDENT IS REQUIRED TO: Identify the "class" the lever belongs to	
	Student OMITS answer	Student gives WRONG answer
EXAMPLES	On another test item with another example belonging to the same INPUT CLASS, i.e., *levers which have the fulcrum between the effort force and the resistance force*	
	-The student also gives no answer	-The student gives the same wrong answer e.g., *classifies it as a third class lever*
	On probes the student does one or more of the following:	
	-Says: e.g., *I don't know* e.g., *I don't know why* e.g., *I don't know how you tell which type it is* -Gives none of the responses in the adjacent column	-Attributes properties to INPUT incorrectly: e.g., *says the "resistance force" is between the effort force and the fulcrum* -Attends to irrelevant properties e.g., *says he gave his answer because of the "downward direction" of the resistance force* -Identifies all the possible answers but doesn't know which are required e.g., *identifies properties of levers correctly and their differences, but doesn't know which label goes with which*
LIKELY PROBLEM	Learning has NOT occurred	Incorrect learning has occurred

F.3.2

DETERMINING THE TYPES OF PROBES TO USE TO DETERMINE SOURCE OF ERROR: PRESENTED IN THE ORDER (i.e., 1-4) IN WHICH THEY SHOULD BE ASKED

DECISION MATRIX TYPE OF ERROR	A Action OMITTED on RECALL test item	B WRONG action taken on RECALL test item	C Action OMITTED on TRANSFER test item	D WRONG action taken on TRANSFER test item
1. USING ADDITIONAL TEST ITEMS	*-Administer another RECALL item* *-If action omitted again, proceed to #2* *-If wrong action taken, proceed as in B*	*-Administer another RECALL item* *-If action omitted, proceed as in A* *-If same or different wrong answer given, proceed to #2 below*	*-Administer another TRANSFER item* *-If action omitted again* *OR* *-If wrong action taken, as in D, administer a RECALL item and proceed as in A or B*	*-Administer another TRANSFER item* *-If action omitted* *-If same or different wrong action taken, administer a RECALL item and proceed as in column A or B*
2. OPEN-ENDED PROBE	*··What's giving you trouble on this question (or problem)?* *··Why did you leave out the answer?*	*··Why did you give the answer you did?*		
3. LESS OPEN-ENDED PROBE	*··Are there any answers you might have given but definitely ruled out?* *··Are there any changes that could be made in the question that would have enabled you to answer the question?*	*··What is there about the _____ (INPUT) that made you give the specific answer you did?* *··What were other possible answers? Why didn't you give them?*		
4. MORE DIRECT PROBE	*··If I call attention to these properties of the _____ (INPUT), can you answer the question?* *··Here are the possible answers; do you know which is the right one?*	*··Which of these properties of the _____ (INPUT) were you considering: ___, ___, ___?* *··Which of them did you consider relevant/ irrelevant?*		

F.3.2

DETERMINING HOW TO USE ADDITIONAL TYPES OF PROBES TO IDENTIFY SPECIFIC, SUSPECTED LEARNING FAILURES

DECISION MATRIX

TYPES OF SUSPECTED PROBLEMS	Incorrect DISCRIMINATIONS	Incorrect GENERALIZATIONS	Incorrect ASSOCIATIONS
ACTION TO TAKE	-*Present additional examples from a different INPUT class which is represented in the original test item and ask:* ••*Are these the same or different?* *OR* ••*Should these be handled or treated the same way or in different ways?* -*If examples vary dimensionally, present pairs of INPUT examples (using the original example as one member) which start far apart and gradually are made more similar* ••*Determine when and where discrimination breaks down*	-*Present additional examples from the same INPUT class which is represented in the original test item and ask:* ••*Are these the same or different?* *OR* ••*Should these be handled or treated the same way or in different ways?* -*If examples vary dimensionally, present pairs of INPUT examples (using the original example as one member) which start close together and gradually are made more dissimilar* ••*Determine when and where generalization breaks down*	-*If behavior has been taught in two directions: INPUT → ACTION and ACTION → INPUT:* ••*Reverse a test question and determine whether same mismatch occurs* -*Identify for the student all the classes of inputs and all the actions that can be taken and then ask:* ••*Which _____ goes with which _____?* (ACTION / INPUT) ••*This is how I pair them up. Am I right or wrong?*

SECTION J of the HANDBOOK provides detailed consideration of ways to interpret student responses to probes. If also deals with interpretation of errors on single and on multiple test items. "Interpretation" is presented at that point in the development process at which it has practical implications; i.e., it provides the basis for determining how to revise instructional materials.

JOB PROCEDURES

	page
TYPES OF INFORMATION YOU WISH TO OBTAIN FROM THE USE OF PROBES	124
DETERMINING WHETHER TO DEVELOP PROBING PROCEDURES	125
Two points during the development process when probes may be selected and/or prepared	126
Information to review during TEST DEVELOPMENT only in order to identify probes which are needed	127
SUMMARY OF PROCEDURES	128
Adequacy of probes	129

F.3.2

CRITERIA FOR IDENTIFYING TYPES OF INFORMATION YOU WISH TO OBTAIN BY USING PROBING QUESTIONS AND PROCEDURES*

IDENTIFICATION MATRIX

TYPES OF INFORMATION TO BE OBTAINED	TYPE OF FAILURE WHICH HAS OCCURRED	VARIABLES ACCOUNTING FOR THE FAILURE
CRITERIA	-Has no learning taken place	-The sources of "no learning occurring" lie primarily in the nature of the instructional materials, for example, in the amount of review provided or in the use of transfer practice items, etc. Probes generally are not designed to identify this type of information. Responses to probes may, on the other hand, identify sections of the instructional materials which have confused students
	-Has incorrect learning taken place involving: ••Discriminations ••Generalizations ••Associations	-Have learning failures occurred due to: ••Student attention to irrelevant properties of INPUTS ••Student confusion created by similarity/dissimilarity of INPUT examples ••Number of properties of INPUTS or of ACTIONS to attend to

*WHILE OTHER DIAGNOSTIC PROCEDURES INVOLVING THE ANALYSIS OF TEST RESULTS CAN LEAD TO AN IDENTIFICATION OF THE TYPE OF LEARNING FAILURE, IT IS ONLY THE PROBING PROCEDURES WHICH DIRECTLY AND EMPIRICALLY IDENTIFY THE REASONS FOR THE FAILURE.

F.3.2

DETERMINING WHETHER TO DEVELOP PROBING PROCEDURES

DECISION MATRIX

CONDITIONS	-Statement of objectives includes subcriterion objectives (indicating probability of learning difficulties) -Associated learning analyses also reveal probable learning difficulties -The number of criterion test items is relatively small (thus making interpretation of test results more difficult)	-Statement of objectives does NOT include subcriterion objectives -Associated learning analyses do NOT reveal probable learning difficulties -The number of criterion test items is relatively large (thus making interpretation of test results easier)
ACTION TO TAKE	*Definitely consider developing and using probing questions and/or procedures*	*Give some consideration to developing and using probing questions and/or procedures**

**The probing procedure is, as already pointed out, the surest direct way of identifying the factors leading to each type of failure.*

F.3.2

CRITERIA FOR IDENTIFYING TWO POINTS IN THE DEVELOPMENT PROCESS WHEN PROBES ARE SELECTED AND/OR PREPARED

IDENTIFICATION MATRIX

TIME IN THE DEVELOPMENT PROCESS	FIRST*	SECOND
CRITERIA	-*During the development of criterion and subcriterion test items, i.e., during TASK F: "DEVELOP DIAGNOSTIC AND EVALUATIVE TESTS"* -*At this point in time, preparation of probes is based on anticipated learning failures*	-*During the evaluation of already prepared instructional materials, i.e., during TASK J: "EVALUATE INSTRUCTIONAL MATERIALS"* -*At this point in time, probes are changed and/or new probes are developed based on test errors which students have actually made in test items*

**Performance of Sub-STEP F.3.2 at this stage of the development process (i.e., during development of TESTS) can be considered "optional." Development of probes can, without loss, be postponed until test results have become available and you wish to probe for sources of specific learning failures you suspect.*

F.3.2

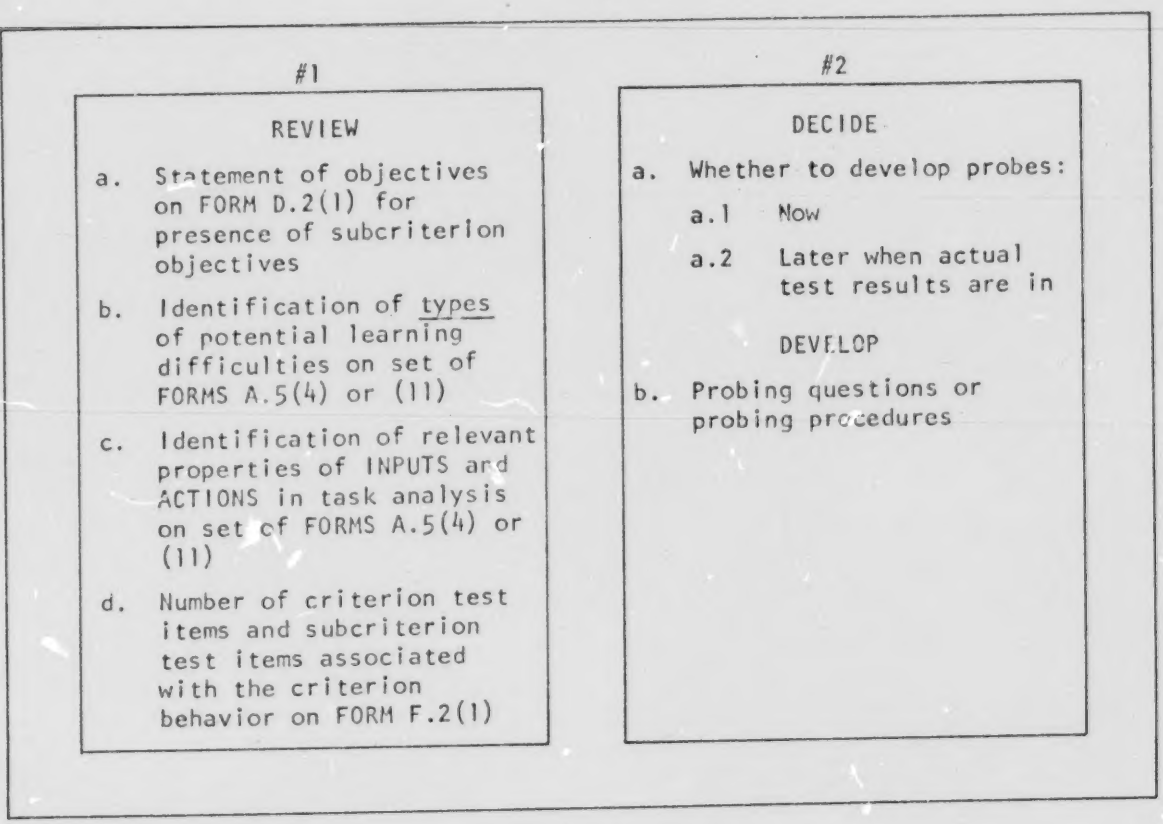
DECISION MATRIX

DETERMINING WHAT INFORMATION TO REVIEW AS A BASIS FOR DEVELOPING PROBES CONCERNING STUDENT ERRORS ON CRITERION AND PREPARATORY TEST ITEMS

INFORMATION NEEDED	Are there likely to be learning failures?	What kinds of information should the probes be designed to get?	Is there likely to be difficulty interpreting results on criterion tests?
ACTION TO TAKE	*REVIEW:* *-Statement of objectives on FORM D.2(1) for presence of* *••Preparatory objectives* *-Collection of all learning analysis forms [FORM A.5(4) or (11)] associated with the criterion behavior for identification of likely learning failures*	*REVIEW:* *-Task analysis forms [FORM A.5(4) or (11)] in order to identify:* *••The criterion inputs and actions* *••Their properties* *-Collection of all learning analysis forms [FORM A.5(4) or (11)] associated with the criterion behavior for identification of:* *••What type(s) of learning failures are likely to occur* *••Potential sources of the difficulties*	*REVIEW:* *-Criterion test items on FORM F.2(1) in order to identify:* *••The number of criterion and subcriterion test items devoted to the criterion behavior*

F.3.2

ILLUSTRATION SUMMARIZING PROCEDURES INVOLVED IN DECIDING WHETHER TO DEVELOP AND IN DEVELOPING DIAGNOSTIC PROBING PROCEDURES


#1
REVIEW
a. Statement of objectives on FORM D.2(1) for presence of subcriterion objectives
b. Identification of types of potential learning difficulties on set of FORMS A.5(4) or (11)
c. Identification of relevant properties of INPUTS and ACTIONS in task analysis on set of FORMS A.5(4) or (11)
d. Number of criterion test items and subcriterion test items associated with the criterion behavior on FORM F.2(1)
#2
DECIDE
a. Whether to develop probes:
a.1 Now
a.2 Later when actual test results are in
DEVELOP
b. Probing questions or probing procedures

F.3.2

PROPERTIES FOR ASSESSING THE ADEQUACY OF THE PROBES WHICH HAVE BEEN DEVELOPED

STANDARDS MATRIX

PROPERTIES	Probes available which are GRADED for amount of CUING provided to student	RELEVANCE	APPROPRIATENESS FOR AUDIENCE
CRITERIA	*-Types of probes which should be available:* *••Fully open-ended* *••Less open-ended* *••Highly directive* *-Planned order of usage:* *••Same order as above*	*-Probes should be geared to the identification of:* *••Suspected or anticipated types of learning failures* *••Factors which are possible reasons for the anticipated failures*	*-Probes should provide tasks which students can easily perform* *-Wording of probes should be:* *••Brief* *••Clear* *••In language audience can understand*

PREVIEW OF THE NEXT SubSTEP

YOUR PRODUCT	*A diagnostic test appropriate to anticipated types of learning failures.*
WHAT YOU WILL WORK FROM	(1) Statements of objectives. (2) Task analysis and learning analysis results identifying likelihood of learning failures. (3) Criterion test items.
WHAT YOU WILL DO	(1) Develop diagnostic test items geared to anticipated types of learning failures.
FORMS YOU WILL USE	FORM F.2(1) for developing diagnostic test items.

DESCRIPTION OF Sub-STEP	F.3.3

INPUT	ACTION	OUTPUT
Statement of objectives + task analysis results + learning analysis results + criterion test items vii	Develop diagnostic test items appropriate to anticipated types of learning failures viii	DIAGNOSTIC TEST ITEM to diagnose type of learning failure ix

Job Aid Contents

CRITERIA FOR IDENTIFYING INPUTS	ACTION TO BE TAKEN	STANDARD FOR OUTPUTS	FORMS TO USE
-MATRIX: Types of diagnostic test items . . . 133-137	-MATRIX: Priorities in order of administering test items . . 140 -MATRIX: Information to review as a basis for developing diagnostic test items . . 141	-MATRIX: Adequacy of diagnostic test items . . . 143	F.2(1) SUMMARY OF PROCEDURES . . . 142

Required Materials

COMPLETED MATERIALS	STEP	COMPLETED FORMS	STEP	BLANK FORMS
Determination of likely need of diagnostic information	F.3.2	Forms associated with criterion behavior (carried forward from	F.3.2	FORM F.2(1)

Sub-STEP F.3.3

JOB DIAGRAM

INPUT	ACTION	OUTPUT
Statement of objectives + task analysis results + learning analysis results + criterion test items vii	Develop diagnostic test items appropriate to anticipated types of learning failures — viii	DIAGNOSTIC TEST ITEM to diagnose type of learning failure — ix

ANTICIPATED PROBLEMS		TEST ITEMS INCLUDE
DISCRIMINATIONS among INPUTS — vii.a	Develop diagnostic test items appropriate to anticipated problem — viii.a	Two or more INPUT examples from same and/or different INPUT classes; testing for ability to see difference between classes of INPUTS* — ix.a
GENERALIZATIONS among INPUTS — vii.b	Develop diagnostic test items appropriate to anticipated problem — viii.b	Two or more INPUT examples from same and/or different INPUT classes: testing for ability to see similarity within a class of INPUTS — ix.b
ASSOCIATIONS between INPUTS and ACTIONS — vii.c	Develop diagnostic test items appropriate to anticipated problem — viii.c	Two or more pairings of INPUTS AND ACTIONS, some of which are correct and some incorrect: testing for ability to identify correct pairings — ix.c
GENERALIZATIONS across ACTIONS — vii.d	Develop diagnostic test items appropriate to anticipated problem — viii.d	Two or more ACTION examples from same and/or different classes: testing for ability to see similarity of ACTIONS within a class of ACTIONS — ix.d

**When it is pairs of ACTIONS that are presented, the student can be tested for his ability to see differences among actions (which may be a problem in some learning situations).*

BACKGROUND INFORMATION

	page
Types of information results on diagnostic test items should reveal	134
Testing for student ability to see "sameness" or "differences"	135
Testing for student ability to see "correctness" or "incorrectness"	135
Types of recognition items useful in diagnosis	136
Developing recognition items for diagnosing specific types of problems	137

F.3.3

CRITERIA FOR IDENTIFYING WHAT IT IS YOU WANT TO FIND OUT BY USING DIAGNOSTIC TEST ITEMS

IDENTIFICATION MATRIX

POTENTIAL PROBLEMS	DISCRIMINATIONS among INPUTS	GENERALIZATIONS across INPUTS	ASSOCIATIONS between INPUTS and ACTIONS	GENERALIZATIONS across ACTIONS
CRITERIA	*Can the student see the difference between:* *-All the INPUT classes* *-Just some of the INPUT classes* *-None of the INPUT classes*	*Can the student see the similarity between:* *-Members belonging to each INPUT class* *··For all classes* *··For some classes* *··For no classes*	*Has the student correctly associated INPUT classes and their relevant ACTIONS for:* *-All the INPUT classes* *-For some INPUT classes* *-For no INPUT classes*	*Can the student see the similarity between:* *-Members belonging to each ACTION class* *··For all classes* *··For some classes* *··For no classes*

F.3.3

DETERMINING HOW TO TEST FOR STUDENT ABILITY TO "SEE" SIMILARITY OR TO "SEE" DIFFERENCES

DECISION MATRIX

TESTING FOR STUDENT ABILITY	To see SIMILARITY	To see DIFFERENCE
ACTION TO TAKE	*-Use RECOGNITION test items* *-When testing for ability to see similarity or differences among INPUTS* *••Do not require as the answer an identification of the action associated with one or more of the inputs* *e.g., do not ask: "Do you say 'IS' for each of these nouns or for both of them?"* *••Do require as the answer a neutral identification of sameness or difference* *e.g., "Are these the same or different?"* *e.g., "Do you treat these the same way or in different ways?"* *-When testing for ability to see similarity or differences among ACTIONS* *••Do not require an identification of inputs* *••Do require (simply) whether actions can or cannot be used interchangeably* *e.g., "Are these the same or different?"* *e.g., "Are these equivalent?"* *e.g., "Would you use these under the same circumstances or different circumstances?"*	

F.3.3

DETERMINING HOW TO ALLOW STUDENT TO INDICATE "CORRECT" VS. "INCORRECT"

DECISION MATRIX

TESTING FOR STUDENT ABILITY	To see CORRECTNESS	To see INCORRECTNESS
ACTION TO TAKE	*-Use RECOGNITION test items* *-When testing for ability to see correctness or incorrectness of associations between INPUTS and ACTIONS:* *••Simply require an indication of correctness vs. incorrectness*	

F.3.3

IDENTIFICATION MATRIX

CRITERIA FOR IDENTIFYING TYPES OF RECOGNITION ITEMS USEFUL IN DIAGNOSING ERRORS

TYPES OF ITEMS	Using PAIRS OF EXAMPLES as the "GIVEN" in a test item	Using MULTIPLE EXAMPLES as the "GIVEN" in a test item
CRITERIA	*Re: INPUTS* *-Pairs of examples from two different classes are used to determine ability to see differences* *••As many pairings of examples from two classes as there are combinations* *e.g., if there are three classes, three pairs would be tested* *e.g., if there are four classes, six pairs would be tested** *-Pairs of examples from the same class to determine the ability to see similarities* *••At least one pair representing each class* *Re: ACTIONS* *-Pairs of examples of actions from different classes* *-Pairs of examples of actions from same class* *Re: INPUTS and ACTIONS* *-Pairs of correct and incorrect associations*	*Re: INPUTS or ACTIONS* *-Lists or collections of examples* *••May all be from same class* *••May all be from different classes* *••May be a mixture* *-Student must identify:* *••Whether examples are the same or different* *OR* *••Which examples are the same or different* *Re: INPUTS AND ACTIONS* *-List or collection of examples of association* *••May be all wrong* *••May be all right* *••May be a mixture* *-Student must identify:* *••Whether they are correct or incorrect associations* *••Which are correct and which are incorrect*
EXAMPLES	*e.g., FOUR CLASSES (A, B, C, D) *The pairings would be as follows:* ••AB ••BC ••AC ••BD ••AD ••CD	

F.3.3 DETERMINING HOW TO DEVELOP RECOGNITION TEST ITEMS TO DIAGNOSE SPECIFIC TYPES OF LEARNING PROBLEMS

DECISION MATRIX

TYPES OF PROBLEMS TO DIAGNOSE	Diagnosing problems in DISCRIMINATION among INPUTS	Diagnosing problems in GENERALIZATION across INPUTS	Diagnosing problems in ASSOCIATING INPUTS and ACTIONS	Diagnosing problems in GENERALIZING across ACTIONS
ACTION TO TAKE	-*Pair INPUTS from different classes and require students to indicate whether they see them as "similar" or as "different"* -*Present multiple examples of INPUTS (from same and different classes) and require students to indicate whether they see them as "similar" or as "different"*	-*Pair INPUTS from the same class and require students to indicate whether they see them as "similar" or as "different"* -*Present multiple examples of INPUTS (from same and different classes) and require students to indicate whether they see them as "similar" or as "different"*	-*Present a correct and an incorrect association of INPUTS and ACTIONS and require students to indicate which is correct* -*Present multiple examples and require students to indicate which is correct and which is incorrect*	-*Pair ACTIONS from the same class and require students to indicate whether they see them as "similar" or as "different"* -*Present multiple examples of ACTIONS (from same and different classes) and require students to indicate whether they see them as "similar" or as "different"*
EXAMPLES	#1 GIVEN: (a) *Two examples (one a perfectly elastic object--a steel spring compressed--and one a non-perfectly elastic object--a squashed ball of steel wool);* (b) *Question:* "Are these the same or different?" NOTE: The question does not identify the action by saying which is perfectly elastic and which is not perfectly elastic (See page 135) #2 GIVEN: (a) *A variety of examples (some perfectly elastic, some not perfectly elastic) are presented;* (b) *Student must group (or check, or point to) those that are similar*	#1 GIVEN: (a) *Two first class levers are paired;* (b) *Question:* "Are these levers the same type or different?" #2 GIVEN: (a) *Multiple examples of lever types for all three classes;* (b) *Question:* "Which of these belong together?"	#1 GIVEN: (a) *Two association examples are used--a lever type labeled "1st class," AND the same lever labeled "3rd class";* (b) *Question:* "Which is the correct labeling?" #2 GIVEN: (a) *Multiple examples of labeling--some correct, some incorrect;* (b) *Question:* "Which of these are correctly labeled?"	#1 GIVEN: (a) *Two examples of actions are paired:* -*Testing for significance of differences between means of testing:* ..t test ..F test (b) *Question:* "Are these "equivalent" or "non-equivalent"? NOTE: By testing to see whether actions are seen as similar or different, it is possible to determine whether, in addition to being able to generalize across actions, whether students can discriminate between actions. This can sometimes be a problem and may be a source of difficulty in associating INPUT and ACTION correctly.

JOB PROCEDURES

	page
Priorities in the order in which diagnostic type tests are administered	140
Information to review in planning the preparation of diagnostic test items	141
SUMMARY OF PROCEDURES	142
Adequacy of diagnostic test items	143

*As in the case of the development of probes, diagnostic test items can be developed at the time other types of test items are being developed. Their development can be postponed until Task J when instructional materials are being tried out and errors on instructional materials and on criterion tests have become available. If these results (and/or results of probing procedures) do not unambiguously identify the specific type of learning failure, diagnostic tests may be developed which are specific to the residual problems in diagnosing learning failures.

F.3.3 DETERMINING THE APPROPRIATE ORDER IN WHICH TO ADMINISTER THE DIFFERENT TYPES OF DIAGNOSTIC TESTS (Either in STEP F.4 or in TASK J)

DECISION MATRIX	1:	2	3
CONDITIONS	No diagnostic tests have yet been administered	Results of first items administered reveal NO difficulties in discriminations or generalizations regarding INPUTS	Results of first items administered reveal NO difficulties regarding ACTION generalization or in discriminations among ACTIONS
ACTION TO TAKE	*FIRST* *-Administer diagnostic items testing for:* *••Discriminations* *••Generalizations regarding INPUTS*	*SECOND* *-Administer diagnostic items testing for:* *••Action generalizations* *••Discriminations among ACTIONS (optional)*	*THIRD** *-Administer diagnostic items testing for:* *••Associations between INPUTS and ACTIONS*

**By postponing testing of ASSOCIATIONS until last, it is possible to make a differential diagnosis of an "association" problem. If administered earlier, items involving both INPUTS and ACTIONS would produce results which are indeterminate as to the type of learning failure.*

F.3.3

DECISION MATRIX

DETERMINING WHAT INFORMATION TO REVIEW AS A BASIS FOR DEVELOPING DIAGNOSTIC TEST ITEMS CONCERNING STUDENT ERRORS ON CRITERION AND PREPARATORY TEST ITEMS

INFORMATION NEEDED	Are there likely to be learning failures?	What kinds of information should the diagnostic items be designed to get?	Is there likely to be difficulty interpreting results on criterion tests?
ACTION TO TAKE	*REVIEW:* *-Statement of objectives on FORM D.2(1) for presence of* *..Suberiterion objectives* *-Collection of all learning analysis forms [FORM A.5(4) or (11)] associated with the criterion behavior for identification of likely learning failures*	*REVIEW:* *-Task analysis forms [FORM A.5(4) or (11)] in order to identify:* *••The criterion inputs and actions* *••Their properties* *-Collection of all learning analysis forms [FORM A.5(4) or (11)] associated with the criterion behavior for identification of:* *••What type(s) of learning failures are likely to occur* *••Potential sources of the difficulties*	*REVIEW:* *-Criterion test item on FORM F.2(1) in order to identify:* *••The number of criterion and suberiterion test items devoted to the criterion behavior*

F.3.3

ILLUSTRATION SUMMARIZING PROCEDURES INVOLVED IN PREPARING DIAGNOSTIC TEST ITEMS

#1

REVIEW

a. Statement of objectives on FORM D.2(1) for presence of subcriterion objectives

b. Identification of types of potential learning difficulties on set of FORMS A.5(4) or (11)

c. Identification of relevant properties of INPUTS and ACTIONS in task analysis on set of FORMS A.5(4) or (11)

d. Number of criterion test items and subcriterion test items associated with the criterion behavior on FORM F.2(1)

#2

DECIDE

a. Whether to develop diagnostic test items:

a.1 Now

a.2 Later when actual test results are in

DEVELOP

b. Diagnostic test questions: Use FORM F.2(1) to record each test item

F.3.3

STANDARDS MATRIX

CRITERIA FOR ASSESSING THE ADEQUACY OF DIAGNOSTIC TEST ITEMS

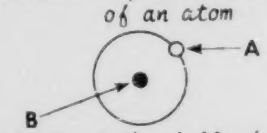
PROPERTIES	RELEVANCE and COVERAGE	AVOIDANCE OF NON-DIFFERENTIAL RESULTS	PLANNED ORDER OF ADMINISTRATION
CRITERIA	*-Test items should be:* *••Fully open-ended* *••Less open-ended* *••Highly directive* *-Planned order of usage:* *••Same order as above* *-Sampling of test items should be sufficient to test for all the anticipated problems*	*-Student responses should NOT require identification of the ACTION associated with an INPUT* *-Student responses should only require identification of* *••Same/different* *••Correct/incorrect*	*-Test items sequenced to test for problems in this order:* *FIRST: Discriminations and generalizations regarding INPUTS* *SECOND: ACTION generalizations** *THIRD: Associations between INPUTS and ACTIONS*

*Where applicable: Discriminations regarding ACTIONS which may be confused

F.3.3

EXAMPLES ILLUSTRATING DIAGNOSTIC TEST ITEMS GEARED TO PROVIDE INFORMATION ABOUT SPECIFIC TYPES OF ANTICIPATED LEARNING FAILURES

EXAMPLES

TYPE OF ANTICIPATED PROBLEM	Re: INPUTS -DISCRIMINATIONS -GENERALIZATIONS	Re: ACTIONS -GENERALIZATIONS -DISCRIMINATIONS (an optional diagnosis)	Re: INPUTS and ACTIONS -ASSOCIATIONS
	e.g., TEST ITEM *There are two problems:* (1) (X) - (Y) (2) (X) + (-Y) *Should they be treated:* ___ *In the same way* ___ *In different ways* Diagnoses a possible "generalization" problem. Failure to generalize across two sets of symbols which signify the same thing. e.g., TEST ITEM *Here are two problems:* (1) (X) - (Y) (2) (X) - (-Y) *Should they be treated:* ___ *In the same way* ___ *In different ways* Diagnoses a possible "discrimination" problem. Failure to discriminate between two sets of symbols which require different treatment.	e.g., TEST ITEM *You wish to measure the magnitude of the buoyant force exerted on a submerged object.* *If you went about measuring it in the following two ways, would you get the same result or a different result?* (1) *Measuring the difference in weight when the object is weighed in air;* AND (2) *Measuring the weight of the overflow that results from submerging the object.* *The measurements that result will be:* ___ *Equivalent* ___ *Non-equivalent*	e.g., TEST ITEM GIVEN: *Representation of an atom*  *Which of the following locates particles in an atom correctly?* ___ (1) *Proton at A, electron at B* ___ (2) *Electron at A, proton at B* Diagnoses possible failure to associate the electron and proton labels (ACTION) with the appropriate location represented (INPUTS)

STEP F.3

COMPLETION CHECKLIST

	IDENTIFIED	PERFORMED	PRODUCED	FORMS COMPLETED
F.3.1	Need for supplementary diagnostic test items	Planned use of criterion test items for diagnostic purposes		
F.3.2			Diagnostic probing questions for interviewing students	
F.3.3			Diagnostic test items	

STEP F.4

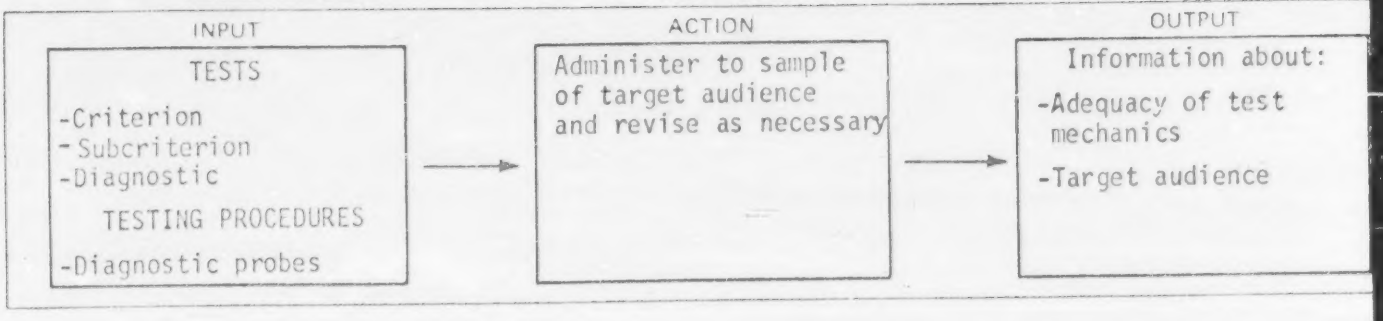
F.4 Try out and revise testing procedures.

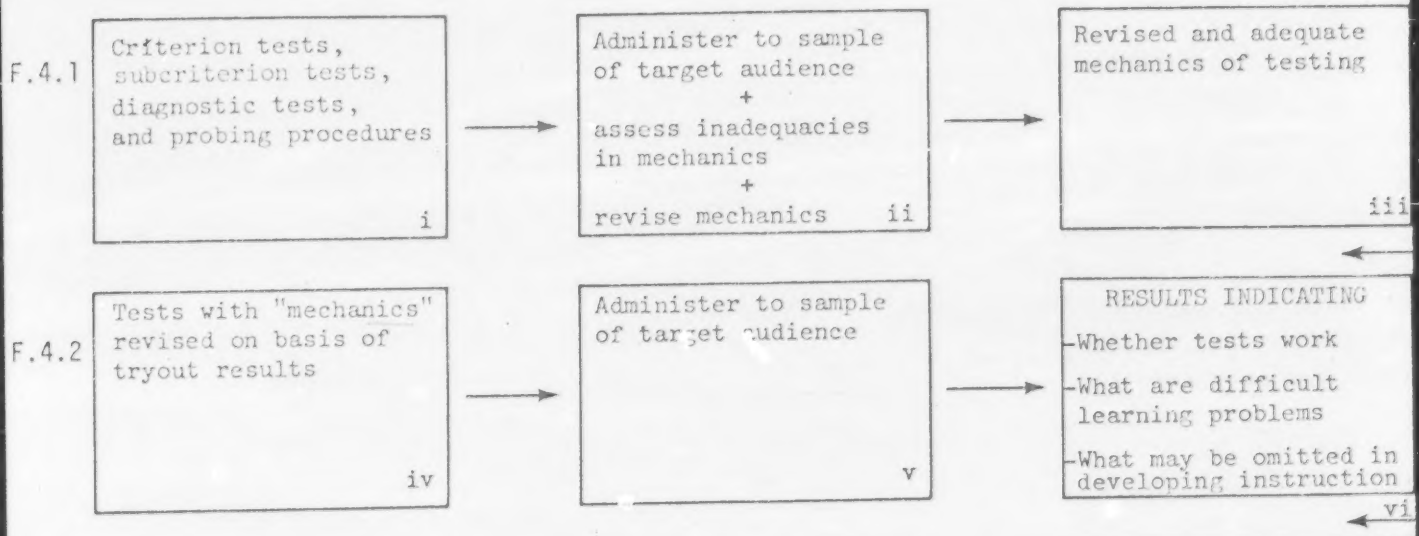
F.4.1 Administer testing procedures to a sample of the target audience as a means of assessing and upgrading the mechanics of the testing procedures.

F.4.2 Perform an analysis of the target audience by administering the revised tests.

STEP F.4

OVERVIEW





STEP F.4

PAGE INDEX

	CRITERIA FOR IDENTIFYING INPUTS	ACTION TO BE TAKEN	STANDARD FOR OUTPUTS	FORMS TO USE
F.4.1	-MATRIX: What features of test mechanics to assess . . 154, 155	-MATRIX: How to assess mechanics of testing . . 156	-MATRIX: Adequacy of assessment of testing procedures . . 159	SUMMARY OF PROCEDURES . 158
F.4.2	-MATRIX: Types of information revealed by tryout results . . 163-165		-MATRIX: Assessing adequacy of test tryouts . 169	SUMMARY OF PROCEDURES . . 168

PREVIEW OF THE NEXT SubSTEP

YOUR PRODUCT	*A revised set of testing procedures which have been empirically tried out.*
WHAT YOU WILL WORK FROM	(1) All tests (criterion, subcriterion, diagnostic). (2) Probing procedures.
WHAT YOU WILL DO	(1) Administer above to a sample of the target audience. (2) Assess inadequacies in mechanics of testing procedures (instructions, statement of problems, recording of results, probing procedures, etc.). (3) Revise mechanics.
FORMS YOU WILL USE	None

DESCRIPTION OF Sub-STEP	F.4.1

INPUT	ACTION	OUTPUT
Criterion tests, subcriterion tests, diagnostic tests, and probing procedures i	Administer to sample of target audience + assess inadequacies in mechanics + revise mechanics ii	Revised and adequate mechanics of testing iii

Job Aid Contents

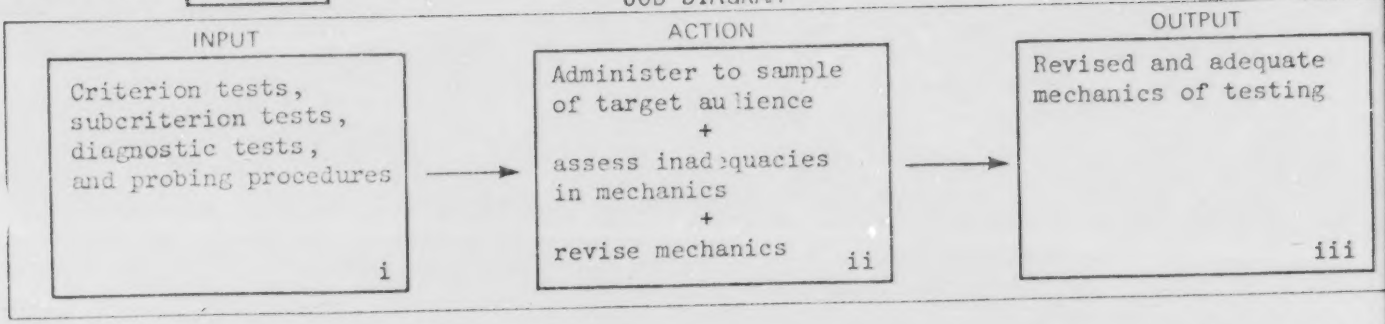
CRITERIA FOR IDENTIFYING INPUTS	ACTION TO BE TAKEN	STANDARD FOR OUTPUTS	FORMS TO USE
-MATRIX: What features of test mechanics to assess . . 154, 155	-MATRIX: How to assess mechanics of testing . . 156	-MATRIX: Adequacy of assessment of testing procedures . . 159	SUMMARY OF PROCEDURES . . 158

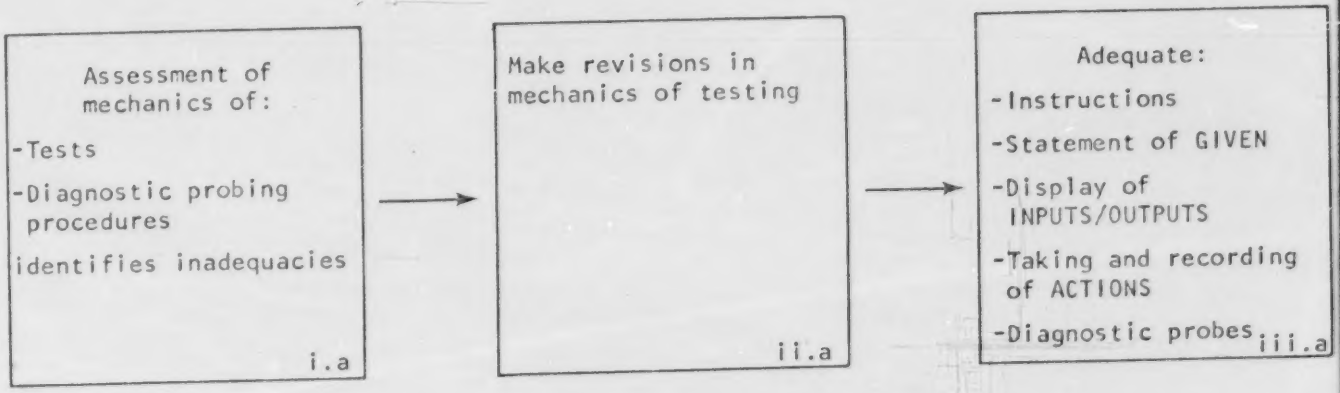
Required Materials

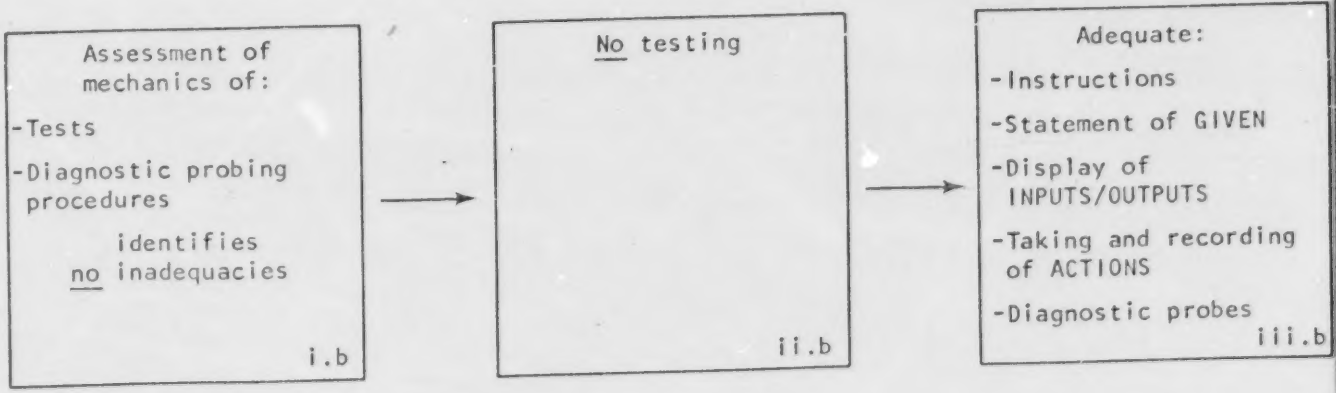
COMPLETED MATERIALS	STEP	COMPLETED FORMS	STEP	BLANK FORMS
Diagnostic probing procedures (carried forward from)	F.3.3	Criterion, subcriterion, and diagnostic tests (carried forward from)	F.3.3	

Sub- STEP F.4.1

JOB DIAGRAM







BACKGROUND INFORMATION

	page
Identification of verbal components of testing procedures which should be assessed	154
Identification of functional properties of testing procedures which should be assessed	155
How to obtain information needed to assess adequacy of testing procedures	156

F.4.1

IDENTIFICATION MATRIX

CRITERIA FOR IDENTIFYING WHAT TYPES OF INFORMATION ARE NEEDED IN ASSESSING THE ADEQUACY OF VERBAL COMPONENTS OF TESTING PROCEDURES

TYPES OF INFORMATION NEEDED	Are INSTRUCTIONS adequately understood	Are statements of TEST QUESTIONS adequately understood
CRITERIA	*Does the student understand:* *-What INPUTS or INPUT properties to attend to or observe* *-What the task is he is expected to perform; what ACTION(s) he is expected to exhibit* *-What he is expected to turn out (OUTPUTS)*	*Does the student understand the language used to describe:* *-The GIVEN in a test question or problem* *••Description of INPUTS* *••Description of ACTION options (in multiple choice tests)* *-The questions asked in probing interviews designed to produce diagnostic information*

SEE PAGE 156 FOR WAYS TO MAKE THIS ASSESSMENT AND THE ASSESSMENT DESCRIBED ON THE OPPOSITE PAGE.

F.4.1

IDENTIFICATION MATRIX

CRITERIA FOR IDENTIFYING TYPES OF INFORMATION NEEDED TO ASSESS THE FUNCTIONAL ADEQUACY OF TESTING PROCEDURES

TYPES OF INFORMATION NEEDED	Are visual and/or verbal INPUTS adequately DISPLAYED	Is there an adequate opportunity for the student to RESPOND, i.e., take an ACTION	Are visual and/or verbal OUTPUTS adequately DISPLAYED*
CRITERIA	*Is the student able adequately to:* *-See or read INPUTS* *-Hear INPUTS* *-Feel, taste, or smell INPUTS*	*Is the student able adequately to:* *-Write or record his ACTION* *-Produce the required motor ACTIONS*	*Is the student able adequately to:* *-See or read OUTPUTS* *-Hear OUTPUTS* *-Feel, taste, or smell OUTPUTS*

**Adequate display of OUTPUTS is particularly important when the OUTPUTS become the INPUTS for serial ACTIONS in a chain.*

	INPUT Properties:	ACTION Properties:	OUTPUT Properties:
EXAMPLES	-Legibility or visibility -Audibility -Manipulability -Capability of being tasted or smelled	-Space or locations to record ACTIONS ••Space to write ••Space to check options ••Access to places to point to -Opportunity to take motor ACTIONS ••Equipment features that work (knobs to turn, switches to operate, etc.)	-Legibility or visibility -Audibility -Manipulability -Capability of being tasted or smelled

F.4.1

DECISION MATRIX

DETERMINING HOW TO OBTAIN INFORMATION NEEDED TO ASSESS THE ADEQUACY OF TESTING PROCEDURES (TWO WAYS WHICH CAN BE USED SEPARATELY OR TOGETHER)

ASSESSMENT CONDITIONS	During student PERFORMANCE on testing procedures	During INTERVIEWS with students held during or following administration of test items or test procedures
ACTION TO TAKE	*-Make observations of what student does to determine if:* *••He attends to relevant INPUTS or INPUT properties (rather than to irrelevant ones)* *••He attempts the required task rather than some other task* *••He produces the required output type rather than something else* *-Make a record of when and over what students:* *••Indicate confusion* *••Spontaneously ask questions* *••Indicate an inability to make out what the INPUTS are* *••Indicate problems or difficulties in taking ACTIONS*	*-Solicit from students information about:* *••Their difficulties or confusions* *••What they thought about instructions or tasks which were involved* *••Whether they can see or read INPUTS* *••Whether they had problems taking any of the required ACTIONS* *••Whether they understood "probing" questions*

JOB PROCEDURES

	page
SUMMARY OF PROCEDURES	158
Standards for adequacy of assessment	159

F.4.1

ILLUSTRATION SUMMARIZING PROCEDURES INVOLVED IN ASSESSING AND REVISING THE MECHANICS OF TESTING PROCEDURES*

#1

a. Administer tests (criterion, subcriterion, or diagnostic) and/or administer diagnostic probing procedures to small sample or target audience:
   - -Those already proficient and/or
   - -Those who are to receive instruction

b. Observe student performance during testing procedures to obtain information about deficiencies in the mechanics of testing procedures

c. Make a record during testing of spontaneous student indications of difficulties

d. Following conclusion of testing procedures, interview students about difficulties

#2

a. Make revisions in testing procedures:
   - -In instructions**
   - -In statement of test problems or questions
   - -In display of INPUTS
   - -In provisions to allow ACTIONS to be taken and/or recorded
   - -In display of resulting OUTPUTS

***See Section G (FORMULATION OF INSTRUCTIONAL STRATEGIES) for guidelines on "instructions" designed to control student attending or observing behavior*

**The tryout and revision of testing procedures should be carried out more than once (when feasible to do so) until students perform as required (e.g., attend to the proper INPUTS or attempt to perform the task required of them).*

F.4.1

CRITERIA FOR EVALUATING THE PROCESS OF ASSESSING AND REVISING THE MECHANICS INVOLVED IN TESTING PROCEDURES

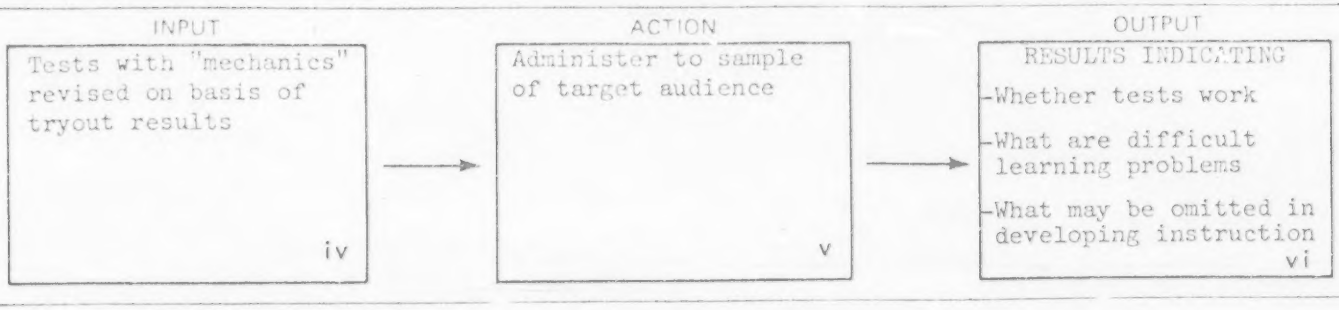
STANDARDS MATRIX

PROPERTIES	COMPLETENESS OF ASSESSMENT/REVISION	VALIDITY OF ASSESSMENT/REVISION	RELIABILITY OF ASSESSMENT/REVISION
CRITERIA	*Covers all of the following:* *-Statement of "GIVEN" in test problem* *-Display of INPUTS* *-Provision for making and/or recording ACTIONS* *-Display of INPUTS*	*-Assessment is based on data:* *Observations* *Interviews* *(Is NOT based on subjective judgments by you, the designer, as to adequacy of mechanics)* *-Assessment is based on data from members of the target audience* *-Revision is a process, continuing through cycles until new assessment shows mechanics to be adequate*	*-Assessment is based on data from a sample (at a minimum, a half dozen people) of the target audience*
	Covers the following when applicable: *-Diagnostic test probes used in interviews*		

PREVIEW OF THE NEXT SubSTEP

YOUR PRODUCT	*An identification of:* *--the adequacy of the test;* *--anticipated learning problems;* *--what may be omitted from the instructional program.*
WHAT YOU WILL WORK FROM	(1) Tests and revised testing procedures.
WHAT YOU WILL DO	(1) Administer tests to a sample of the target audience (2) Review results as a basis for determining adequacy of tests and making decisions about the development of the program.
FORMS YOU WILL USE	None

DESCRIPTION OF Sub-STEP	F.4.2



Job Aid Contents

CRITERIA FOR IDENTIFYING INPUTS	ACTION TO BE TAKEN	STANDARD FOR OUTPUTS	FORMS TO USE
-MATRIX: Types of information revealed by tryout results . . 163-165		-MATRIX: Assessing adequacy of test tryouts . 169	SUMMARY OF PROCEDURES . . 168

Required Materials

COMPLETED MATERIALS	STEP	COMPLETED FORMS	STEP	BLANK FORMS
		Tests with mechanics revised based on tryout results	F.4.1	

Sub-STEP F.4.2

JOB DIAGRAM

INPUT	ACTION	OUTPUT
Tests with "mechanics" revised on basis of tryout results iv	Administer to sample of target audience v	RESULTS INDICATING -Whether tests work -What are difficult learning problems -What may be omitted in developing instruction vi
Results of tests administered to sample of BEGINNERS iv.a	Analyze results v.a	Identification of -Whether BEGINNERS can already exhibit criterion behavior -Whether they have previously learned behaviors which can interfere with the learning of criterion behavior vi.a
Results of tests administered to sample of students who have COMPLETED existing instruction iv.b	Analyze results v.b	-Whether COMPLETERS have residual difficulties with respect to the criterion behavior vi.b
Difference in results of tests administered to COMPLETERS and BEGINNERS iv.c	Analyze results v.c	-Whether the test tests what it is supposed to vi.c

BACKGROUND INFORMATION

	page
Identifying potential samples from a target audience for tryout of test items	164
Identifying types of information (and its uses) which tryout of tests provides	165

F.4.2

IDENTIFICATION MATRIX

CRITERIA FOR IDENTIFYING TWO TYPES OF SAMPLES FROM THE SAME TARGET AUDIENCE

TYPES OF SAMPLES	Those who have already COMPLETED instruction on criterion behaviors	Those who are BEGINNERS
CRITERIA	-*Those who have just satisfactorily completed an existing course or training program* -*Those who have been on a job (for some time) and are judged to be performing adequately*	-*Those who are about to begin a course or training program* -*Personnel just about ready to begin a job*
EXAMPLES	e.g., *students who have just completed a portion of or a particular topic in a third grade math course*	e.g., *students who are about to begin a portion of or a particular topic in grade math course*
	Students in both groups are roughly matched for general intelligence or ability	
	e.g., *teachers on the job*	e.g., *beginning teacher trainees*

F.4.2

CRITERIA FOR IDENTIFYING TYPES OF INFORMATION TO BE DERIVED FROM TRYOUT RESULTS ON TESTS WITH TWO TYPES OF SAMPLE FROM TARGET AUDIENCE

IDENTIFICATION MATRIX

TYPES OF TEST RESULTS	Results for BEGINNERS on tests: ••Criterion tests ••Subcriterion tests	Results for COURSE COMPLETERS* on tests: ••Criterion tests ••Subcriterion tests ••Diagnostic tests ••Diagnostic probing procedures	DIFFERENCE in results for COURSE COMPLETERS and BEGINNERS on tests: ••Criterion tests ••Subcriterion tests
CRITERIA	*Results identify:* *-Whether beginners already have learned:* *••The criterion behavior* *••Parts of the criterion behavior* *-Whether the beginners have previously learned:* *••Behaviors which may interfere with the learning of the criterion behavior*	*Results identify:* *-Whether completers have residual difficulties with:* *••The criterion behavior* *••Parts of it* *••Component skills*	*Results identify:* *-Whether the test adequately tests for proficiency at:* *••The criterion behavior* *••Parts of it*
HOW THE INFORMATION CAN BE USED	-Can influence the decision of whether or not to develop instructional materials for a particular criterion behavior (or part of it) ••If all beginners can exhibit a particular criterion behavior, no instruction on it need be prepared -Can influence the decision of whether or not to allow some students to bypass instructional materials covering that criterion behavior ••If some students can exhibit the criterion behavior and some cannot, in accommodating individual differences, you may plan to allow those who can to bypass the related instructional materials (resulting in more efficient use of instructionaltime) -Can influence formulation of instructional strategies to overcome potential interference existing behaviors may have with the learning of new ones e.g., old associations which may interfere with the learning of new ones	-Can feed into the strategy formulation (See Section G) so that instruction on the weak area can be directed to specific learning problems **Completion of existing courses, not the one you are planning to develop.*	-Can be used to revise test items which do not adequately measure what they are intended to measure e.g., paper-and-pencil items which are badly constructed may lead to no test differences, even when beginners have not previously learned the criterion behavior

JOB PROCEDURES

	page
SUMMARY OF PROCEDURES	168
Standards for adequacy of tryout	169

F.4.2

CRITERIA FOR ASSESSING THE ADEQUACY OF TRYOUT OF TESTS AND TEST PROCEDURES

STANDARDS MATRIX

PROPERTIES	SAMPLING OF TARGET AUDIENCE	COMPLETENESS OF ANALYSIS OF TEST RESULTS
CRITERIA	-*Target audience represented by both:* ••*Beginners* ••*Completers (of previous courses or programs)* -*Each type of sample includes a minimum of ten people*	-*Analysis attempts to identify:* ••*Residual difficulties completers have (as a basis for focusing on the problems during formulation of instructional strategies)* ••*What beginners already can do (as a basis for individualizing scheduling)* ••*What beginners already do (as a basis for determining what may interfere with what has to be learned)*

STEP F.4

COMPLETION CHECKLIST

	IDENTIFIED	PERFORMED	PRODUCED	FORMS COMPLETED
F.4.1		Assessed adequacy of mechanics of testing procedures	A revised set of tests and/or diagnostic probing procedures	
F.4.2	-Potential learning problems -What target audience (beginners) already know	Tried out tests with sample of target audience	A revised set of tests	